

**The Characterisation of a Novel**  
**Deubiquitinating Enzyme in *Escherichia coli***

Debora Ruth Channing, MSc.

Thesis submitted to the University of Nottingham  
for the Degree of Doctor of Philosophy

JULY 2010

## Abstract

Although bacteria do not contain ubiquitin or ubiquitin homologues, the accurate proteolytic processing of ubiquitin precursors and ubiquitin fusion proteins in lab strains of *Escherichia coli* has previously been noted<sup>3, 4</sup>. We provide evidence that a novel ubiquitin-fusion processing activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) represents a specific DUB activity against linear (peptide-linked) ubiquitin fusions. Fusions of ubiquitin linked to an ATP-binding cassette protein (LmrC) or to enhanced green fluorescent protein (EGFP), expressed in Rosetta<sup>TM</sup>2(DE3) were cleaved precisely after the C-terminal Gly76 of ubiquitin. The use of gene knock-out showed that the source of the ubiquitin-fusion processing activity in Rosetta<sup>TM</sup>2(DE3) is the ubiquitin-like protease, *elaD*; as specific ubiquitin-fusion processing was ablated by the inactivation of the *elaD* gene. Whilst this study was in progress, Catic *et al.* showed that *elaD* is present in the commensal *E. coli* strain K12 and intestinal pathogenic strains, but absent from extraintestinal pathogenic strains, and exhibits deubiquitinating activity *in vitro* against the generic substrate ubiquitin-AMC<sup>14</sup>. Our study has demonstrated that *elaD* not only exhibits deubiquitinating activity against linear ubiquitin fusions, but also possesses isopeptidase activity, with a preference for unanchored Lys63-linked poly-ubiquitin chains over Lys48-linked forms. GST-*elaD* has also been shown in this study to bind specifically to

immobilised mammalian ubiquitin in pull-down assays. Thus, elaD is a bacterial enzyme which has the ability to functionally interact with the highly conserved eukaryotic ubiquitin protein and indicate that elaD may have a role in regulating host-microbe interactions.

## **Acknowledgements**

I wish to thank Dr Robert Layfield and Dr Ian Kerr, my supervisors for their help and guidance during the course of this work. I am grateful also to the Biotechnology and Biological Sciences Research Council for funding this project.

My thanks are extended to Barry Shaw, Dr James Cavey and Jo Strachan for their expertise and support in teaching me some of the molecular biological techniques.

I would also like to acknowledge my partner Ray for his patience and support in ensuring that the use of the computer and information technology ran as smoothly as possible whilst writing my thesis. My thanks also go to my daughter, Victoria for kindly locking up the chickens, ducks and turkeys when I worked late in the Winter times.

# Contents

<b>ABSTRACT .....</b>	<b>2</b>
<b>1 INTRODUCTION.....</b>	<b>12</b>
INTRODUCTION.....	12
1.2 COMPARISON OF THE EUKARYOTIC-UBIQUITIN MEDIATED SYSTEM WITH PROKARYOTIC SYSTEMS.....	12
1.2.1 Ubiquitin and UBLs.....	14
1.2.2 The eukaryotic ubiquitin-mediated system.....	16
1,2,2,1 Four areas of protein targeting in a ubiquitin-mediated system .....	16
1,2,2,2 The ubiquitin-mediated system in signalling pathways .....	20
1.2.3 The signals which identify the target proteins .....	27
1.2.4 Ubiquitination of a target protein .....	31
1.2.5 Deubiquitination of a target protein .....	37
1.2.5.1 The catalytic mechanisms of proteases.....	37
1.2.5.2 Tag removal as part of the Eukaryotic ubiquitin-mediated system.....	41
1.2.5.3 Tag removal for virulence .....	46
1.2.5.4 Substrate specificity of DUBs or ULPs .....	47
1.2.6 Disposal of a target protein .....	51
1.3 MICROBES WHICH SUBVERT HOST CELL SYSTEMS .....	54
1.4 DUB ACTIVITY OBSERVED IN NON PATHOGENIC E. COLI .....	60
1.5 THE MAIN AIMS FOR THIS STUDY .....	62
<b>2 MATERIALS AND METHODS.....</b>	<b>64</b>
2.1 CELLS, MEDIA AND PREPARATION OF LYSATES.....	64
2.1.1 Cells and media .....	64
2.1.2 Preparation of cell lysates.....	65

2.2 OLIGONUCLEOTIDES USED IN THE STUDY AND STANDARD NUCLEIC ACID TECHNIQUES.....	66
2.2.1 <i>Oligonucleotides used in this study</i> .....	66
2.2.2 <i>Plasmid preparation</i> .....	66
2.2.3 <i>Obtaining genomic DNA to use as a DNA template in PCR</i> .....	66
2.2.4 <i>Polymerase chain reaction (PCR) and site directed mutagenesis (SDM)</i> .....	66
2.2.4.1 PCR using Pfu DNA polymerase (Promega) .....	66
2.2.4.2 PCR using GoTaq <sup>®</sup> DNA Polymerase (Promega).....	67
2.2.4.3 SDM.....	67
2.2.4.3.1 SDM using Pfu DNA polymerase (Promega): .....	67
2.2.5 <i>Further treatment of amplified DNA products:</i> .....	67
2.2.5.1 PCR products for cloning:.....	67
2.2.5.2 PCR products for SDM or to generate <i>elaD</i> knock-outs: .....	68
2.2.6 <i>Screening for colonies using colony PCR</i> .....	68
2.2.7 <i>Restriction endonuclease digestion</i> .....	68
2.2.8 <i>DNA precipitation for use in generating <i>elaD</i> knock-out in K12</i> .....	69
2.2.9 <i>Ligation</i> .....	69
2.2.10 <i>Phage 1 (P1) transduction</i> .....	69
2.2.11 <i>DNA electrophoresis</i> .....	70
2.3 TRANSFORMATION.....	70
2.3.1 <i>Transformation of chemically competent bacteria</i> .....	70
2.3.1.2. Transformation of competent bacteria.....	70
2.3.2 <i>Transformation of bacteria using electroporation in generating the <i>elaD</i> gene</i>	
<i>knock-out JIG182<math>\Delta</math><i>elaD</i><sup>Apr</sup> used ultimately to make Rosetta<sup>TM</sup> 2(DE3)<math>\Delta</math><i>elaD</i><sup>Apr</sup></i> .....	71
2.3.2.1 Preparation of bacteria for electroporation to generate JIG182 $\Delta$ <i>elaD</i> <sup>Apr</sup> : .....	71
2.3.2.2 Electroporation of K12 to make JIG182 $\Delta$ <i>elaD</i> <sup>Apr</sup> : .....	71
2.3.2.3 Induction of $\lambda$ Red proteins involved in generating JIG182 $\Delta$ <i>elaD</i> <sup>Apr</sup> : .....	72
2.4 STANDARD PROTEIN TECHNIQUES .....	72

2.4.1 Affinity purification of (His) <sub>6</sub> -tagged recombinant proteins.....	72
2.4.2 Affinity purification of GST-tagged recombinant proteins.....	73
2.4.2.1 Capture of GST-tagged proteins on Glutathione-Sepharose™ 4B beads .....	73
2.4.2.2 Eluting GST-tagged proteins from Glutathione-Sepharose™ 4B beads.....	74
2.4.2.3 Thrombin cleavage to remove the GST-tag from GST-tagged proteins.....	74
2.4.3 Dialysis of purified proteins.....	74
2.4.4 Generation of ZNF-Sepharose 4B beads .....	74
2.4.4 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) .....	75
2.4.5 Coomassie staining and destaining .....	75
2.4.6 Transfer of proteins to nitrocellulose membrane for N-terminal sequencing and for Western blotting (WB) .....	76
2.4.6.1 Preparation of proteins for N-terminal sequencing (Edman degradation).....	76
2.4.6.2 WB for (His) <sub>6</sub> -tagged recombinant proteins.....	77
2.4.6.3 WB for ubiquitin.....	77
2.4.6.4 WB for GST-tagged recombinant proteins .....	78
2.4.6.5 WB for SUMO1.....	78
2.5 DEUBIQUITINATION ASSAYS .....	78
2.6 CLONING METHODS AND GENERATING MUTANTS .....	79
2.6.1 Cloning recombinant (His) <sub>6</sub> -NEDD8-LmrC, (His) <sub>6</sub> -SUMO1-LmrC or (His) <sub>6</sub> -Ub-EGFP using a modified pHUE vector .....	79
2.6.2 Cloning recombinant EF-Tu (GST-TufA and GST-TufB) and GST-elaD expression vectors.....	80
2.6.2.1 Cloning TufA and TufB into pGEX4T1.....	80
2.6.2.2 Cloning elaD into pGEX4T1.....	81
2.6.3 SDM to generate (His) <sub>6</sub> -Ub-LmrC mutants: GA, GP and GGP.....	82
2.6.4 SDM to generate GST-elaD mutants: C313S, N227A, W232A, D169A.....	82

**Table 2.3** PCR amplification conditions used to amplify DNA using either Pfu DNA polymerase (Promega) or GoTaq<sup>®</sup> DNA Polymerase (Promega). PCR mixes were then purified using QIAquick PCR Purification Kit (QIAGEN) or gel purified using QIAquick Gel Extraction kit (QIAGEN)..... 85

<b>3 CHARACTERISATION OF A NOVEL UBIQUITIN-FUSION PROCESSING ACTIVITY IN THE <i>E. COLI</i> SUBSTRAIN ROSETTA™2(DE3)</b>	<b>89</b>
3.1 INTRODUCTION	89
3.2 WORK LEADING UP TO THIS STUDY	92
3.3 THE AIMS AND OBJECTIVES OF THE STUDY	95
3.4 RESULTS	96
3.4.1 <i>In vivo</i> cleavage of the (His) <sub>6</sub> -Ub-LmrC fusion protein in <i>E. coli</i> substrain Rosetta™2(DE3)	96
3.4.2 Investigating the nature of the cleavage of the (His) <sub>6</sub> -Ub-LmrC fusion protein	99
3.4.3 Characterising the specificity of the DUB activity in <i>E. coli</i> substrain Rosetta™2(DE3) by modifying residues around the cleavage site in the (His) <sub>6</sub> -Ub-LmrC substrate	101
3.4.4 Edman degradation to ‘pinpoint’ the cleavage site of the GA,	107
3.4.5 Mass spectrometry to determine the precise mass of (His) <sub>6</sub> -Ub and the proteins present in the ~11 kDa region	109
3.4.6 Further characterising of the DUB activity in <i>E. coli</i> substrain Rosetta™2(DE3) by replacing either one of the fusion proteins within the (His) <sub>6</sub> -Ub-LmrC fusion protein...	115
3.4.6.1 Construction of (His) <sub>6</sub> -NEDD8-LmrC, (His) <sub>6</sub> -SUMO1-LmrC and (His) <sub>6</sub> -Ub-EGFP to investigate specificity of the DUB activity in <i>E. coli</i> substrain Rosetta™2(DE3)	117
3.4.6.2 Expression of fusion proteins, (His) <sub>6</sub> -SUMO1-LmrC, (His) <sub>6</sub> -NEDD8-LmrC and (His) <sub>6</sub> -Ub-EGFP in <i>E. coli</i> substrain Rosetta™2(DE3)	120
3.4.7 Investigating the possibility that DUB activity in <i>E. coli</i> was species specific	123



3.5 DISCUSSION .....	125
<b>4 A CANDIDATE APPROACH TO IDENTIFY THE PROKARYOTIC ENZYME RESPONSIBLE FOR THE DUB ACTIVITY IN <i>E. COLI</i> SUBSTRAIN ROSETTA™2(DE3).....</b>	<b>131</b>
4.1 INTRODUCTION .....	131
4.2 RESULTS.....	134
4.2.1 <i>The cloning of GST-tagged candidates EF-Tu and elaD, for the DUB activity in E. coli substrain Rosetta™2(DE3) .....</i>	<i>134</i>
4.2.2 <i>Over-expression of DUB Candidate proteins: GST-TufA, GST-TufB, GST-elaD and eukaryotic elongation factors, GST-EF-1<math>\alpha</math>1, GST-EF1-<math>\alpha</math>2.....</i>	<i>137</i>
4.2.3 <i>Improving purification and yield of the DUB candidate, GST-elaD protein.....</i>	<i>140</i>
4.2.3.1 <i>Increasing the induced yield of DUB candidate, GST-elaD protein .....</i>	<i>140</i>
4.2.3.2 <i>Removing GroEL from purified GST-elaD protein with the aim to produce homogenous pure protein .....</i>	<i>142</i>
4.2.4 <i>An assay to screen the DUB candidates for the ability to bind with ubiquitin .....</i>	<i>144</i>
4.2.2 <i>Generation of the elaD knock-out, Rosetta™2(DE3)<math>\Delta</math>elaD<sup>Apr</sup> to investigate further the possibility that elaD may be responsible for the DUB activity observed in this E. coli substrain.....</i>	<i>147</i>
4.2.3 <i>Expression of (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP recombinant proteins in Rosetta™2(DE3)<math>\Delta</math>elaD<sup>Apr</sup> to serve as substrates for the DUB activity .....</i>	<i>154</i>
4.3 DISCUSSION .....	159
<b>5 FURTHER CHARACTERISATION OF ELAD DUB ACTIVITY .....</b>	<b>163</b>
<b>5.1 INTRODUCTION.....</b>	<b>163</b>
5.2 RESULTS.....	170
5.2.1 <i>An in vitro DUB assay using (His)<sub>6</sub>-Ub-LmrC as a substrate for GST-elaD .....</i>	<i>170</i>

5.2.2 An <i>in vitro</i> DUB assay using unanchored linear di-ubiquitin as a substrate for GST- <i>elaD</i> .....	174
5.2.3 K63 and K48 isopeptide bonds within unanchored tetra-ubiquitin as substrates for GST- <i>elaD</i> .....	178
5.2.4 Construction of a GST- <i>elaD</i> catalytic mutant, GST-C313S for comparison with GST- <i>elaD</i> in <i>in vitro</i> DUB assays.....	182
5.2.4.1 Designing the GST- <i>elaD</i> catalytic mutant, GST-C313S.....	182
5.2.5 K63 and K48 linked unanchored poly-ubiquitin mixed length chains as substrates for immobilised GST- <i>elaD</i> including GST-C313S.....	185
5.2.6 K63 and K48 linked unanchored poly-ubiquitin mixed length chains as substrates for GST- <i>elaD</i> in solution.....	188
5.2.7 Analysis of elongation factors, both prokaryotic (EF-Tu) and eukaryotic (EF-1 $\alpha$ ) for DUB activity using unanchored K63/K48-linked mixed length poly-ubiquitin as substrates.....	190
5.2.8 Anchored poly-ubiquitin chains as substrates for GST- <i>elaD</i> , including exclusive K6, K11, K27, K29, K33, K48, K63 linkages .....	193
5.2.9 A mixture of target proteins, poly-ubiquitinated and mono-ubiquitinated, including unanchored poly-ubiquitin from rat muscle as substrates for GST- <i>elaD</i> .....	196
5.2.10 Characterising GST- <i>elaD</i> DUB activity using unanchored K63-linked mixed length poly-ubiquitin.....	198
5.2.11 Construction of mutations within GST- <i>elaD</i> designed to impair binding of GST- <i>elaD</i> to ubiquitin: GST-N227A, GST-W232A and GST-D169A.....	201
5.2.12 Using mutants of GST- <i>elaD</i> (N227A, W232A and D169A) to investigate recognition of ubiquitin.....	205
5.3 CONCLUSIONS .....	208
<b>6 DISCUSSION .....</b>	<b>214</b>

6.1. KEY FINDINGS OF THIS STUDY .....	214
6.1.1 <i>E. coli</i> substrain Rosetta <sup>TM</sup> 2(DE3) expresses a specific DUB with a conventional mechanism .....	214
6.1.2 The DUB in <i>E. coli</i> substrain Rosetta <sup>TM</sup> 2(DE3) was shown to prefer ubiquitin over SUMO1 or NEDD8 as substrates in vivo .....	215
6.1.3 The DUB in <i>E. coli</i> substrain Rosetta <sup>TM</sup> 2(DE3) was identified as <i>elaD</i> .....	218
6.1.4 GST- <i>elaD</i> prefers unanchored Lysine 63-linked poly-ubiquitin as a substrate in vitro .....	220
6.1.5 GST- <i>elaD</i> -ubiquitin binding and <i>elaD</i> structural neighbours.....	222
6.1.6 Implications and outcomes .....	223
6.2. FUTURE WORK .....	228
6.2.1 Investigating the possible role or roles of <i>elaD</i> .....	228
6.2.2 Further investigations in to the catalytic activity and interaction partners for <i>elaD</i> .....	230
<b>7 REFERENCES .....</b>	<b>232</b>

# **1 Introduction**

## ***Introduction***

The focus of this chapter is to introduce key topics, appropriate to this study as well as the main aims and objectives of this work. The areas encompassed include protein processing systems in eukaryotes and prokaryotes, modes of attack by pathogenic eubacteria and the findings in the literature of deubiquitinating enzyme (DUB) activities in non pathogenic *E. coli*.

## ***1.2 Comparison of the eukaryotic-ubiquitin mediated system with prokaryotic systems***

Ubiquitin is a small (76 amino acid residues), highly conserved protein which is covalently linked to target proteins for proteolysis or to alter the activity of the target protein. Ubiquitin is expressed as precursor ubiquitin either as repeat units of ubiquitin or as fusion proteins to ribosomal units. Cells need systems by which proteins are targeted for their disposal either to maintain a balance between protein translation and proteolysis as well as for protein quality control. Additionally cells need to control the activities of key proteins for control over signalling pathways. Therefore the proteins of signalling pathways are more likely to be targeted to change their interactions with other proteins rather

than for proteolysis. There are many different cellular systems which serve these protein degradation and signalling functions, however in this chapter attention will be restricted to the main system used in eukaryotes, the ubiquitin-mediated system. When used in this text, the ubiquitin-mediated system is used to include signalling systems, protein transport and protein degradation, which use ubiquitin. The area covered will include ubiquitin-like proteins (UBLs). More particularly the focus will be on the components of the eukaryotic ubiquitin-mediated system compared with what is known of prokaryotic systems with related functions.

Three scientists, Aaron Ciechanover, Avram Hershko and Irwin Rose, were crucial to the discovery of ubiquitin-mediated proteolysis and consequently were awarded the Nobel Prize in Chemistry for 2004<sup>23, 24</sup>. The small eukaryotic ubiquitin protein is now known to play a key role in many of the cell systems for example: signalling pathways, quality control of proteins, the cell cycle and cell differentiation. This involvement of ubiquitin in such diverse cell systems means that many disease states involve altered function of the ubiquitin-mediated system. Consequently the ubiquitin-mediated system and its target proteins are an area of intensive investigation, with the aim of using the components of the ubiquitin-mediated system as potential therapeutic targets. An increasing amount of information is being revealed of the ubiquitin-

mediated system and disease states and is covered in a recent review by T. Jung<sup>25</sup>. The next section will begin with the components of the eukaryotic ubiquitin-mediated system and their associated functions will be introduced briefly. This will then be followed by sections detailing each component of the eukaryotic ubiquitin-mediated system, compared with any prokaryotic systems or proteins with a similar structure or function.

### **1.2.1 Ubiquitin and UBLs**

Ubiquitin is found in all eukaryotes and an archaeobacterium, *Thermoplasma acidophilum* expresses ubiquitin as well as a 20S proteasome<sup>26,27</sup>. However, to date there is no functional evidence to describe ubiquitin tagging in the archaea. Curiously, one group also describes the purification of ubiquitin from the eubacterium *Anabena variabilis* which contradicts the current belief that eubacteria lack the gene for ubiquitin<sup>28</sup>. Interestingly, recent findings have revealed that the eubacterium, *Mycobacterium tuberculosis* has a ubiquitin analogue, prokaryotic ubiquitin like protein (Pup) which is 64 amino acid residues compared with ubiquitin which is 76 amino acid residues<sup>29</sup>. Pup is used for protein modification in a similar way to ubiquitin but is thought to use a different mechanism for attachment to the target protein compared with ubiquitin<sup>30-32</sup>.

Ubiquitin and UBLs have a ubiquitin superfold and a flexible glycine, glycine C-terminus<sup>33</sup>. Ubiquitin and UBLs are involved in a variety of eukaryotic cell signalling processes rather than just as part of a protein quality control system. The main UBLs are small Ubiquitin-related Modifier (SUMO) in humans; interferon stimulated gene 15 (ISG15) in humans; neural precursor cell expressed developmentally down-regulated protein 8 (NEDD8) in humans, autophagy associated protein (Atg8) in *Saccharomyces cerevisiae* and ubiquitin-related modifier 1 (Urm1) in *Saccharomyces cerevisiae*. All UBLs, apart from Atg8 are synthesised as inactive precursors, comparable to ubiquitin; therefore they have to be specifically cleaved by an enzyme. For ubiquitin, this is a DUB and for UBLs the enzymes are referred to as UBL specific proteases (ULPs). SUMO is involved in nuclear localisation and transcriptional regulation<sup>34</sup>, as well as in the regulation of circadian rhythms<sup>35</sup>. ISG15 could be involved in transcription and pre-mRNA splicing during IFN response<sup>36, 37</sup>. NEDD8 is associated with transcriptional regulation<sup>18, 38</sup>; Atg8 is involved in autophagy and nutrient recycling in yeast<sup>39, 40</sup> and Urm1 is reported to be involved in budding, nutrient sensing and the oxidative stress response in *Saccharomyces cerevisiae*<sup>41, 42</sup>.

## **1.2.2 The eukaryotic ubiquitin-mediated system**

### **1,2,2,1 Four areas of protein targeting in a ubiquitin-mediated system**

A simplified version of the eukaryotic ubiquitin-mediated system and key components are shown in Fig. 1.1. Many chaperone proteins and facilitative proteins are also involved in the eukaryotic ubiquitin-mediated system, however only the key proteins have been included. This is a system which targets proteins as part of a signalling mechanism or for disposal as part of a protein quality control or as part of autophagy. In all of these cases a protein is ubiquitinated with a particular type of ubiquitin tag, deubiquitinated or disposed of.

The targeting of proteins within the eukaryotic ubiquitin-mediated system can be broken down into four main areas, first, the targeting of the protein for ubiquitination. This involves particular signals which a protein carries singling it out as a target for ubiquitination (not shown in Fig.1.1). The second main area of the eukaryotic ubiquitin-mediated system is the ubiquitination of the target protein. This involves three (sometimes four) enzymes which have a role in the eventual ubiquitination of a target protein; The activation enzyme (E1); conjugation enzyme (E2); ligation enzyme (E3) and in some cases another ligation enzyme, E4. There is a hierarchy of these enzymes as there are only two types of E1s, over thirty types of E2s and hundred of

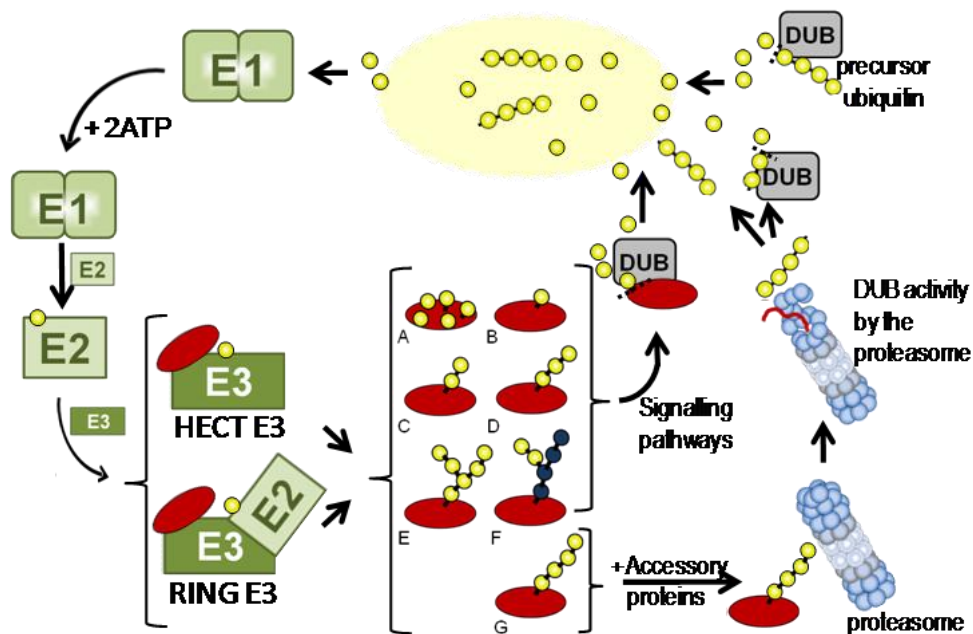


different E3s which adds to the selectivity of these enzymes for their target proteins<sup>43-45</sup>. E1, a homodimer, uses ATP to activate ubiquitin and passes the ubiquitin to E2, which then works with E3 (and sometimes E4) to ubiquitinate a target protein. E3 ligases can be of two main types, each with a slightly different method of ubiquitination; the HECT (homologous to E6 carboxyl terminus) E3 and the RING (really interesting new gene) -type E3. A HECT E3 receives ubiquitin from an E2 and subsequently ubiquitinates the target protein. Alternatively, a RING E3 serves more as a scaffolding protein while the E2 ubiquitinates the target protein. There are many different E2 and HECT or RING-type E3 enzymes which lead to a variety of types of ubiquitination (Fig.1.1 A–G). As will be seen later, UBLs can also be ubiquitinated (Fig.1.1 F).

The third area of the eukaryotic ubiquitin-mediated system is the deubiquitination by DUBs for removing ubiquitin tags, editing ubiquitin tags or cleaving ubiquitin from ubiquitin precursors. Ubiquitin is expressed either as a peptide-linked ubiquitin chain precursor or fused to a carboxyl extension protein (CEP) for example the human HUBCEPs, CEP80 and CEP52<sup>46-48</sup>. This area that has only recently become the focus in studies due to the realisation that DUBs could make excellent therapeutic targets due to their ability to alter the activities of target proteins<sup>49</sup>. Within the fourth area of the eukaryotic

ubiquitin-mediated system is the disposal of target proteins by the proteasome, a complex structure consisting mainly of proteases. Target proteins are recognised by their ubiquitin tag, which the proteasome cleaves for recycling and the target protein is processed.

## The eukaryotic ubiquitin-mediated system



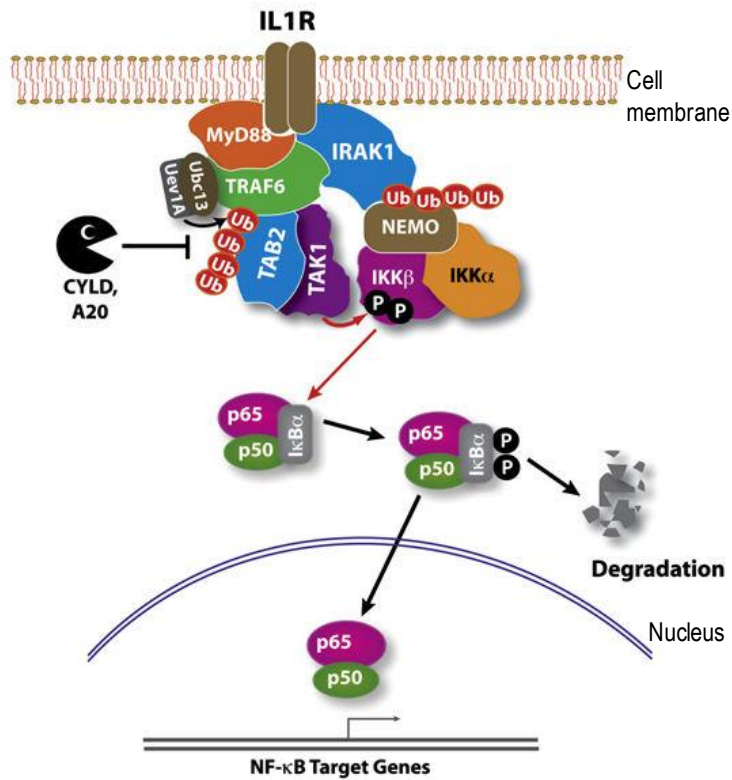
**Fig. 1.1** A cartoon to show the main components of the eukaryotic ubiquitin-mediated system with ubiquitin shown as a yellow sphere, target proteins shown as red ovals and cleaved proteins shown by a dashed line. First the activating enzyme (E1) activates ubiquitin using ATP and passes the ubiquitin to a conjugation enzyme (E2), which in turn works with a ligation enzyme (E3) to ubiquitinate a target protein. There are two types of E3 ligases, the HECT (homologous to E6 carboxyl terminus) E3 and the RING (really interesting new gene) -type E3. If the E2 works with a HECT E3, then the E2 passes ubiquitin to the E3 and the target protein is ubiquitinated by the E3. However, if the E2 works with a RING-type E3 then it is the E2 which ubiquitinates the target protein. There are mainly different E2 and HECT or RING-type E3 enzymes which leads to a variety of types of ubiquitination (A–G), including ubiquitinated ubiquitin like proteins (UBLs), dark blue circles (F). This variety of types of ubiquitination adds to the specificity in the targeting of proteins for modulation of cell signalling pathways. The ubiquitin tag on proteins for disposal will then be recognised by the proteasome, a self compartmentalised protease which deubiquitinates the ubiquitin tag, then processes the target protein. The ubiquitin tags liberated by the proteasome will then form part of the ubiquitin pool as a chain or is cleaved by a DUB and enters the ubiquitin pool as monomers. Alternatively, DUBs also remove ubiquitin tags from ubiquitin precursors or target proteins from signalling pathways to be recycled.

### 1,2,2,2 The ubiquitin-mediated system in signalling pathways

Proteins can be targeted for activation or for the activation of a neighbouring protein or deactivation by the eukaryotic ubiquitin-mediated system as part of a control over cell signalling or in protein transport. The tumour necrosis factor (TNF) receptor associated factor six (TRAF6) is an example of a protein which once ubiquitinated leads to activation of a neighbouring protein within a complex. The inhibitory protein  $\kappa$ B (I $\kappa$ B) is an example of a protein which once ubiquitinated becomes inactivated, leading to the activation of the immune response.

TRAF6 and I $\kappa$ B are targeted as a control over the nuclear factor  $\kappa$ B (NF- $\kappa$ B) signalling pathway, in this case the interleukin-NF- $\kappa$ B signalling pathway. Fig. 1.2 is a schematic view of the interleukin-NF- $\kappa$ B signalling pathway taken from a review by Chen<sup>11</sup>. NF- $\kappa$ B is a transcription factor which is inactive when bound to an inhibitory protein, I $\kappa$ B and is found in the cytosol. When pathogenic microbes are detected by the host cell surveillance systems, as part of an immune response, pro-inflammatory cytokines (for example interleukins) bind to toll like receptors (TLRs). This triggers a chain of events leading to many key proteins forming complexes with the TLRs (in the cytosol). TRAF6 is an E3 which binds to this protein complex and becomes self-ubiquitinated with a poly-ubiquitin chain. This poly-ubiquitin chain binds and activates a kinase within a neighbouring protein complex, the transforming growth factor

(TGF)- $\beta$  activated kinase known as the Tak1 kinase complex. The Tak1 kinase complex is made up of TAK1 binding protein 2 (TAB2) and Tak1 kinase. The poly-ubiquitin chain binds to TAB2 which leads to the activation of Tak1 kinase which phosphorylates the inhibitor of  $\kappa$ B kinase  $\beta$  (IKK $\beta$ ). This then leads to the targeting of the second example of a target protein (I $\kappa$ B) mediated by ubiquitination but this time resulting in inactivating the protein. Phosphorylated IKK $\beta$  adds to the complex and eventually IKK phosphorylates I $\kappa$ B which is then targeted for ubiquitination and consequently I $\kappa$ B undergoes proteolysis in the proteasome. The disposal of I $\kappa$ B leaves NF- $\kappa$ B free to move to the nucleus resulting eventually in the expression of inflammatory proteins used to attack pathogens as part of the immune response.



**Fig.1.2** A schematic view of the interleukin-NF-κB signalling pathway taken from a review by Z.J. Chen and L.J. Sun, to illustrate how ubiquitination and deubiquitination exerts controls over a signalling pathway<sup>11</sup>. In this case the NF-κB proteins are p65 and p50. The NF-κB signalling pathway is activated by pro-inflammatory cytokines released as part of the immune response. A protein complex forms with the TLR. TRAF6, an E3 becomes self-ubiquitinated with a poly-ubiquitin chain. It is this poly-ubiquitin chain that binds to the TAK1 kinase complex leading to the activation of TAK1 kinase which in turn phosphorylates IKKβ. After IKKβ becomes incorporated within the protein complex, IKK then phosphorylates IκBα which is bound to NF-κB. Phosphorylation of IκBα targets it for ubiquitination and consequently inactivation as it is then disposed of by the proteasome. The loss of IκBα allows NF-κB to enter the nucleus and to take part in the transcription of genes involved in the immune response for example inflammatory proteins.

UBLs are also modulators of the activities of targeted proteins in the same way as ubiquitin, sometimes with similar or opposing roles to ubiquitin. For example, NEDD8 increases the ubiquitinating activity of a RING E3, cullin1 (Cul1)<sup>50, 51</sup>. Cul1 forms a protein complex within the, SKP1-CUL1-F-box protein (SCF) E3 ubiquitin ligase complex, E3-SCF<sup>βTrCP</sup> within the interleukin-NF-κb pathway. This increased E3 ligase activity results in the ubiquitination of IκB leaving NF-κB to move to the nucleus and transcribe genes associated with inflammatory protein expression.

NEDD8 and SUMO-1 are found in a signalling pathway which responds to DNA damaged proteins, hypoxia or abnormal proteins by activating a tumour suppressor, p53. Activated p53 increases transcription of genes associated with cell cycle arrest or apoptosis. This protects the body from the proliferation of cells containing mutations, thus preventing tumour formation. In a normal cell p53 is tagged with ubiquitin by a RING E3, Mdm2 and the E2, Ubc5 to keep p53 inactive or at low levels in the cell. There are six lysine residues on p53 available for ubiquitination, K370, K372, K373, K381, K382 and K386. At low levels of Mdm2 in the cell p53 is mono-ubiquitinated, keeping the tumour suppressor inactive. When Mdm2 is at high levels in the cell p53 is poly-ubiquitinated and undergoes proteolysis by the proteasome (Fig.1.3 A).

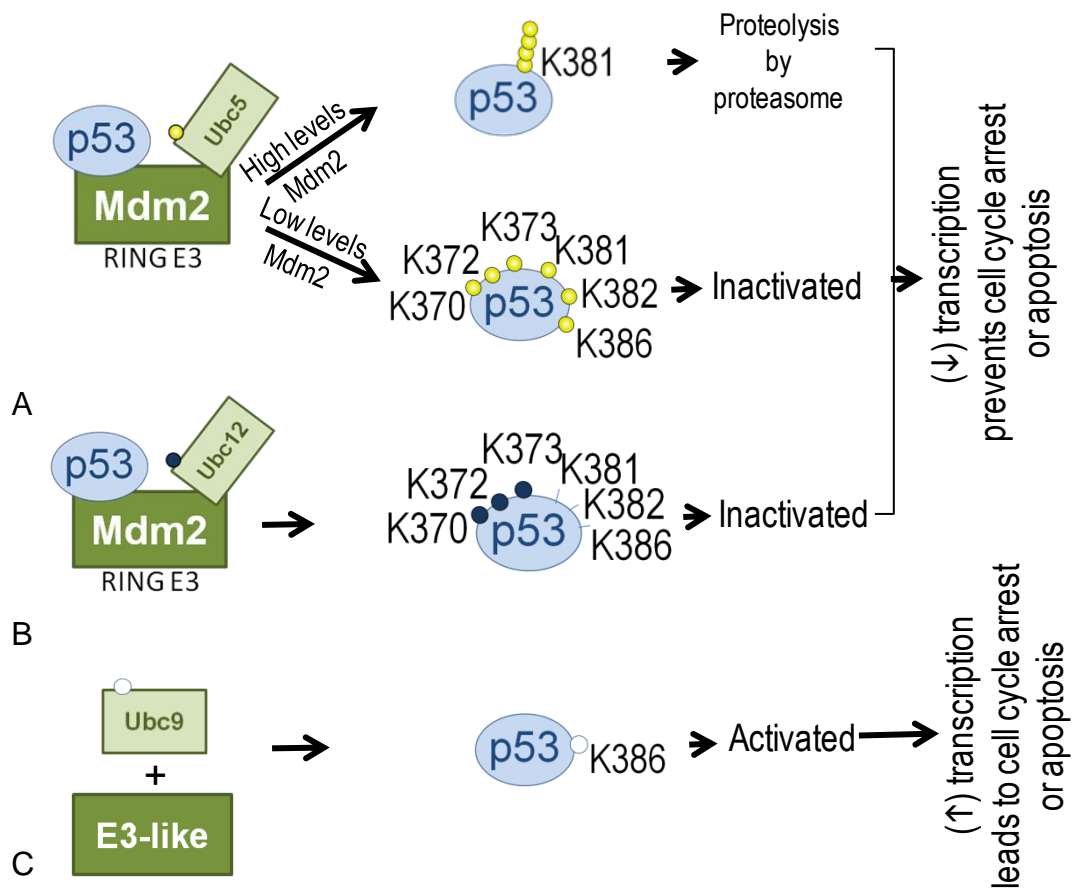
As part of an extra control over the pathway Mdm2 can also self-ubiquitinate and be directed to the proteasome for proteolysis<sup>93</sup>.

NEDD8 affects the activity of p53 within the p53 signalling pathway with the same result as ubiquitinating p53. Mdm2 can self-NEDDylate and NEDDylate p53, this time using E2, Ubc12 (Fig.1.3 B)<sup>18</sup>. The NEDD8 tag on p53 has the same result as a ubiquitin tag, as they both result in the inactivation of p53. There is limited information regarding the reason p53 has an extra controlling tag (NEDD8) to modulate activity as well as ubiquitin.

The SUMOylation of p53 was first demonstrated by M. Gostissa *et al* but this time with the opposite effect of both ubiquitin and NEDD8, resulting in activated p53 (Fig.1.3 C)<sup>22</sup>. Mdm2 was not the SUMOylating enzyme, instead the E2, hUbc9 was responsible for the ligation of SUMO-1 to p53 in the absence of an E3. Other researchers, L. Chen and J. Chen reported that the tumour suppressor, a protein, transcribed from an alternate reading frame of the INK4a/ARF locus (ARF) regulates p53 SUMOylation<sup>52</sup>. ARF forms a complex with Mdm2 and p53, inhibiting the E3 ligase activity of Mdm2, consequently stabilising p53. The ARF-Mdm2-p53 complex is then relocated from the cytosol to the nucleolus where p53 is SUMOylated. The details of how p53 is SUMOylated and the final effects have yet to be ascertained. There



have been reports from other groups which did not see the increased activity of p53 with SUMOylation<sup>53, 54</sup>. However, this was thought to have arisen from different assay conditions<sup>52</sup>.



**Fig.1.3** A schematic view to show the effects of tagging the tumour suppressor, p53 with ubiquitin, NEDD8 or SUMO-1. In a normal cell p53 is continuously ubiquitinated on the six available lysine (K) residues shown in (A), by an E3, Mdm2. The type of ubiquitination tag depends upon the levels of Mdm2, as indicated but both tag types result in the inhibition of p53 transcription activity. This prevents the cell from going in to cell cycle arrest or entering apoptosis<sup>13</sup>. NEDD8 has also been demonstrated to modify p53 at K370, K372 and K373 and the effect is the same as ubiquitin, to keep p53 inactive<sup>18</sup>. It is possible that the three sites which are not modified by NEDD8 can be ubiquitinated resulting in p53 having a two types of tags at the same time. However, there is limited information regarding the significance of p53 also having a NEDD8 modification. SUMO-1 has been shown to tag p53 with the opposite effect of ubiquitin and NEDD8 tags resulting in the cell entering either cell cycle arrest or apoptosis<sup>22</sup>. There is limited knowledge of the details associated with SUMOylation of p53 and the outcomes.

### 1.2.3 The signals which identify the target proteins

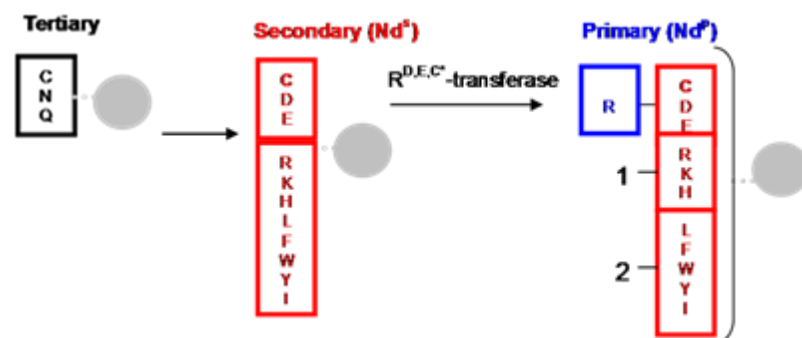
There are three main types of signals which single out proteins to be targeted for signalling, transport or disposal: a sequence in the primary structure or a domain, a polyphosphate (polyP) tag or an SsrA tag.

The first type of signal, a protein sequence within the primary structure, or a domain, is recognised in eukaryotes by an E3, which ubiquitinates the protein, targeting it either for proteolysis by the proteasome or for modification as part of cell signalling. One example of this is a system called the N-end rule, which targets misfolded proteins for proteolysis, first described by A. Bachmair, D. Finley and A. Varshavsky<sup>55, 56</sup>. A polypeptide expressed with a degron at the N-terminus is referred to as having a secondary destabilising N-terminal residue (Nd<sup>s</sup>). In some situations the Nd<sup>s</sup> requires another amino acid residue added by a transferase. This amino acid is termed a primary destabilising N-terminal residue (Nd<sup>p</sup>). In eukaryotes addition of a Nd<sup>p</sup> is carried out by *ATE1*-encoded arginyl-transferase (R<sup>D,E,C\*</sup>-transferase) or by *aat*-encoded Leu/phe-transferase (L/F<sup>K,R</sup>-transferase) in prokaryotes<sup>5, 57-59</sup>. Fig. 1.4 summarises the main degrons which target proteins for proteolysis in mammals and *E. coli*. After ubiquitination, N-recognin then binds directly to the 19S region of the 26S proteasome and the targeted protein undergoes proteolysis. This process was first described in yeast<sup>60</sup>. Another example of a degron in eukaryotes, is the domain

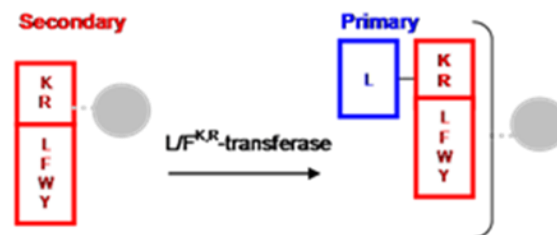
rich in proline (P), glutamate (E), serine (S) and threonine (T), the (PEST) domain, found in I $\kappa$ B, important for degradation by  $\mu$ -calpain rather than the proteasome<sup>61</sup>. There is also the phosphorylation dependent degron, DSGXXS, where X = any amino acid (found in I $\kappa$ B and  $\beta$ -catenin). Both serines in DSGXXS are phosphorylated targeting the protein for interaction with  $\beta$ TrCP of SCF <sup>$\beta$ TrCP</sup>, the protein complex responsible for ubiquitinating proteins targeted for proteolysis<sup>62</sup>. The destruction box (D-box), RXXL (X = any amino acid) is a degron recognised by either APC/C-Cdc20 or APC/C-Cdh1 for ubiquitination and degradation by the proteasome<sup>63</sup>.

In *E. coli* an E3-like protein, ClpS was originally thought to be essential for targeting proteins in the N-end rule pathway for proteolysis by ClpAP in a way similar to that of E3 recognins<sup>1, 64</sup>. However later research revealed that ClpS was not essential for degradation by ClpAP, as ClpAP recognises N-end rule substrates. Instead ClpS was shown to be important for modulating proteolysis as ClpS raises the rate of proteolysis of N-end rule substrates and can also inhibit this proteolysis<sup>57</sup>.

## Mammals



## *E. coli*



**Fig.1.4** N-terminal degradation signals in eukaryotes (mammals) and prokaryotes (*E. coli*). A protein is translated with a secondary destabilising N-terminal residue (Nd<sup>S</sup>), shown in red. In mammals some tertiary signals (shown in black) first have to be modified or converted to an Nd<sup>S</sup>. In some occasions another amino acid residue, R is added by a transferase to form a primary destabilising N-terminal residue (Nd<sup>P</sup>), shown in blue. Red ovals represent target proteins and coloured boxes contain the amino acid degradation signal. Numbers 1 and 2 refer to the substrate binding site within N-recognins which bind to the N-degrons shown<sup>5-10</sup>. Cysteine (C), Asparagine (N), Glutamine (Q), Aspartate (D), Glutamate (E), Arginine (R), Lysine (K), Histidine (H), Leucine (L), Phenylalanine (F), Tryptophan (W), Tyrosine (Y), Isoleucine (I).

In *E. coli* the second type of signal is the polyP tag used for signalling and to target proteins for disposal. The polyP tag is a linear polymer of many hundreds of phosphate residues which is attached to target proteins by polyP (PPKs). The first reported polyP tags were formed part of a stress response in *E. coli* but now it is generally known that polyP tags are also important for their growth and survival<sup>65</sup>. In *E. coli* the polyP tag becomes attached to a target protein labelling it for proteolysis by the Lon protease in response to starvation<sup>66, 67</sup>. In eukaryotes there are many polyP tagged proteins yet little is known of their function<sup>68</sup>. However, there is evidence that polyP plays a role in cell growth and proliferation in mammals by targeting a key enzyme, mammalian target of rapomycin (mTOR) in the mTOR signalling pathway<sup>69-71</sup>. Insulin and amino acids activate mTOR within the mTOR pathway in mammalian cells to initiate translation of genes vital for cell growth and proliferation. The PPK within this pathway is thought to tag mTOR with polyP which activates mTOR.

The third signal to target proteins in this case for disposal, is only found in prokaryotes and involves the co-translational tagging of the C-terminus of a target polypeptide (SsrA), marking it for proteolysis<sup>72</sup>. The main proteases in *E. coli*, involved in the recognition of the SsrA tag are ClpXP and ClpAP, with ClpXP being responsible for the majority (90%) of the degradation of SsrA tagged proteins<sup>73</sup>. This occurs by a process

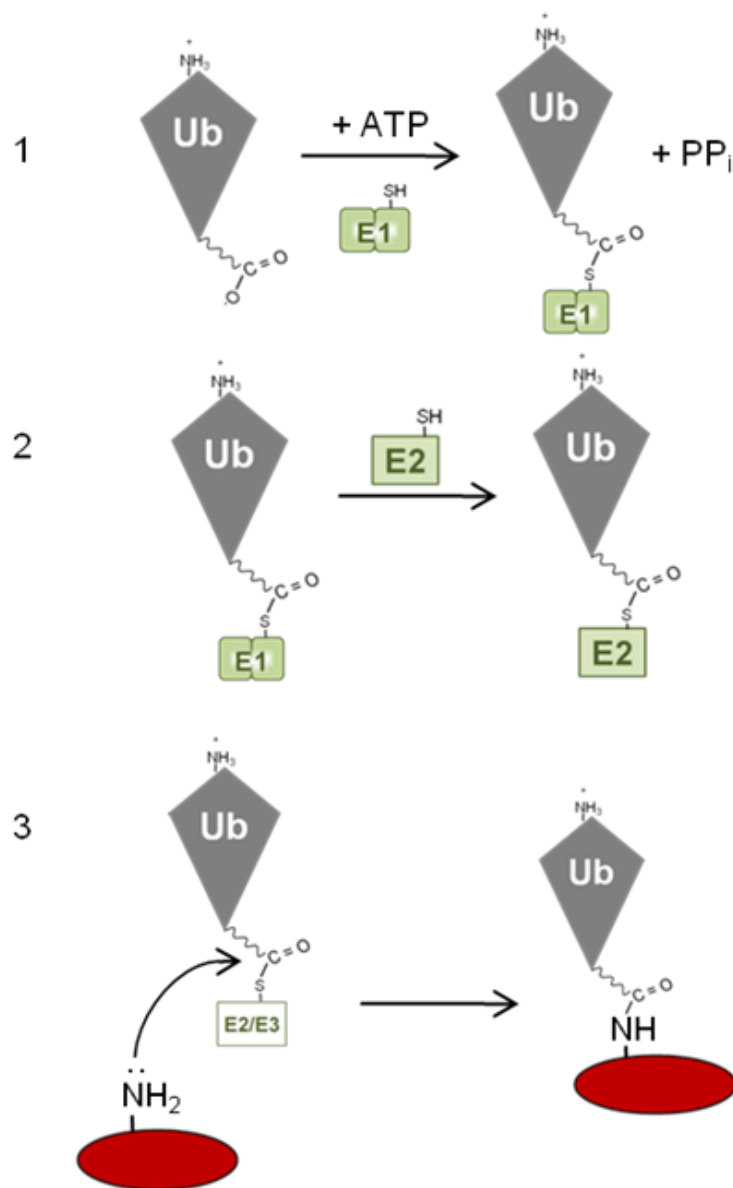
referred to as *trans*-translation. During translation a line of ribosomes move sequentially along the same mRNA, and if a ribosome stalls it needs to be released quickly or it would cause a major obstruction and consequently cessation of translation. The ribosome is rescued by a transfer messenger ribonucleic acid (tmRNA) which is a dual function RNA. This tmRNA is a tRNA charged with alanine (Ala), so is termed tmRNA<sup>ala</sup>, and it carries a mRNA sequence coding for a 10 amino acid residue peptide (SsrA). The ribosome has two binding sites for tRNA molecules: the peptide site (P) for the tRNA bound to the stationary polypeptide-mRNA complex and the acceptor site (A) for incoming tRNA. Charged tmRNA<sup>ala</sup> moves in to binding site A, providing alanine, which causes the mRNA to dissociate from the static polypeptide. The ribosome is then free to translate the sequence carried by the mRNA resulting in an SsrA tag at the C-terminus of the protein, labelling it for subsequent proteolysis. A ribosome can stall either because the mRNA lacks a stop codon due to damage, or when it reaches a rare codon within a complete mRNA<sup>74, 75</sup>.

#### **1.2.4 Ubiquitination of a target protein**

As shown previously the conjugation of ubiquitin to a target protein involves three and sometimes four different enzymes in three steps (Fig.1.5). In step 1, E1 catalyses two subsequent biochemical

reactions which eventually result in the hydroxyl group on the carbon of the  $\alpha$ -carboxyl group of ubiquitin being replaced with a better leaving group, a thiol (SH). First, E1 activates the carbonyl carbon by offering ATP in order to undergo nucleophilic attack by the oxygen (reduced) from the hydroxyl group of the C-terminus of ubiquitin to form a phosphoryl ester. This eventually results in the covalent attachment of the  $\alpha$ -carboxyl group of ubiquitin to a sulfhydryl group in the  $\epsilon$ -amino group of a cysteine residue in the active site of E1. In the second step ubiquitin is transferred to the active site of a second enzyme, E2 which is again bound via a sulfhydryl group of cysteine. In the third step ubiquitin, whilst covalently linked to E2 or E3, undergoes a nucleophilic attack by the  $\epsilon$ -amino group in a specific lysine residue of the target protein. Activation, conjugation and ligation reactions for UBLs occur in a comparable way with those involving ubiquitin using, E1-like, E2-like and E3-like enzymes. An E3 ligase, PafA has been predicted to target proteins with Pup, in the eubacterium, *Mycobacterium tuberculosis* but no other protein tagging system has been found in eubacteria. Therefore it is of interest to note that homologues of E3 ligases have been found to be expressed by pathogenic eubacteria and their function was involved in overthrowing the host cell systems. This will be discussed in section 1.4.



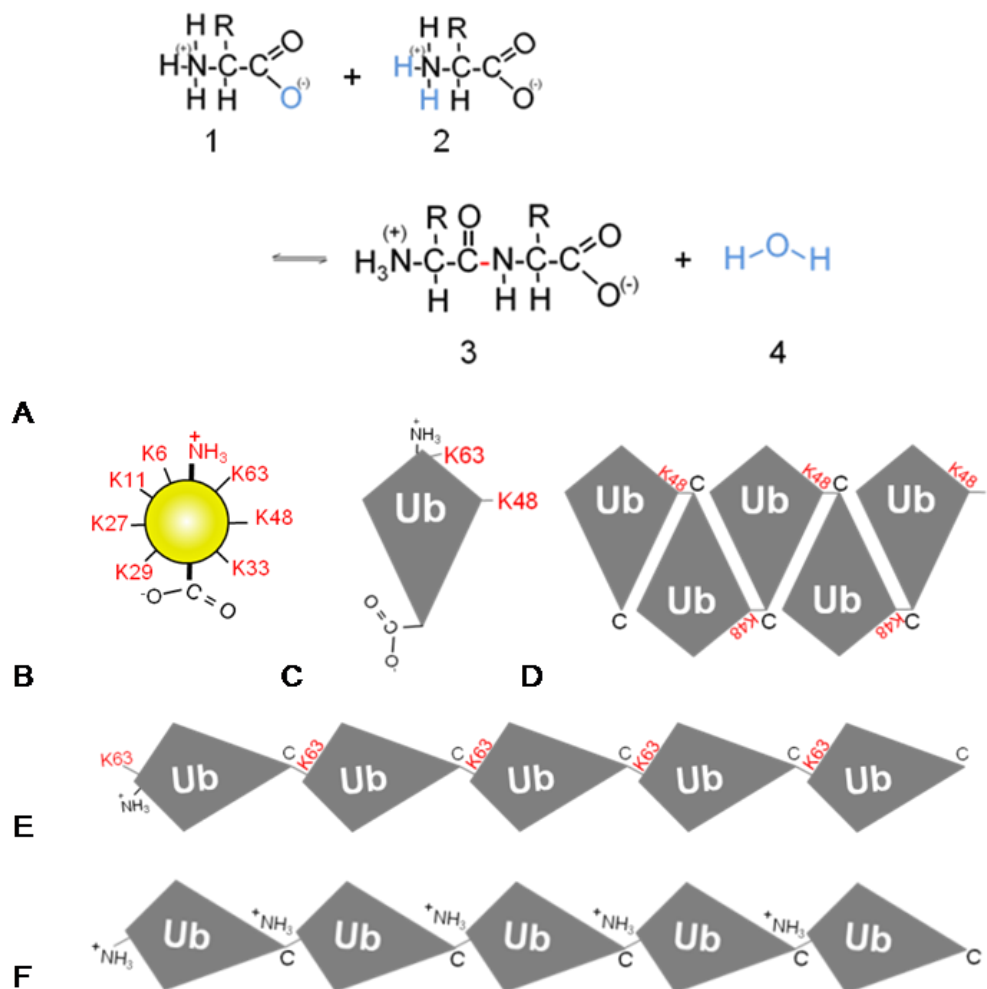


**Fig. 1.5** A cartoon to show the 3 steps in ubiquitinating a target protein (red oval) with ubiquitin (Ub) as a grey kite shape. In step 1, E1 first activates ubiquitin using ATP then the thiol group of cysteine in E1 carries out a nucleophilic attack on the carbonyl carbon of ubiquitin. In step 2, A thiol group of cysteine from an E2 then carries out a nucleophilic attack on ubiquitin. In step 3 ubiquitin, whilst covalently attached to E3 or E2 undergoes nucleophilic attack, this time from a lysine residue within the  $\epsilon$ -amino group of the target protein.

Poly-ubiquitin chain formation sometimes involving a poly-ubiquitin chain conjugation factor (E4) can either be built by adding another ubiquitin protein (donor) to a lysine residue on a ubiquitin (acceptor) of the ubiquitin chain. This is a condensation reaction, shown in Fig.1.6 A. This could involve the  $\alpha$ -amino group or the  $\varepsilon$ -amino group from one of seven possible internal residues in the ubiquitin acceptor: K6, K11, K27, K29, K33, K48 and K63 demonstrated *in vitro and in vivo*<sup>76, 77</sup>. Fig.1.6 B shows ubiquitin as a sphere with all available 7 lysine residues and the  $\alpha$ -amino group. Ubiquitin is also shown as a kite shape in Fig.1.6 C-F. If the  $\alpha$ -amino group of the ubiquitin acceptor is used in poly-ubiquitin chain building this forms a linear poly-ubiquitin chain with an open conformation (Fig.1.6 F). A linear peptide bond exists between precursor ubiquitin proteins or is formed post translationally by an E3, linear ubiquitin chain assembly complex (LUBAC)<sup>78</sup>. When the  $\varepsilon$ -amino group of the ubiquitin acceptor is used this can result in various poly-ubiquitin chain isomers or even in chains with mixed linkages.

Poly-ubiquitin chains connected using the same lysine residue on each subsequent ubiquitin molecule, for example all k48-linked, or all K63-linked in poly-ubiquitin chains (Fig.1.5 D and E respectively) have different conformations from one another. A K63-linked poly-ubiquitin chain is more open compared with a K48-linked poly-ubiquitin chain.

However, recent findings have shown that K<sup>48</sup>-linked poly-ubiquitin chains vary with pH and adopts a fully closed conformation at a pH 7.5 or greater and fully open when pH is 5.4 or less<sup>16</sup>. Ubiquitin chains can also be formed with mixed linkages, for example K48 and K29 can be used to form a forked poly-ubiquitin chain<sup>79</sup>. SUMO chains which have been ubiquitinated can also occur<sup>80, 81</sup>.



**Fig.1.6** The formation of a peptide bond is a condensation reaction between two amino acid residues (1) and (2), with leaving groups shown in blue, to form a peptide bond (red) in the building peptide chain (3) and water (A). Ubiquitin is shown as a yellow sphere with all 7 lysine (K) residues (with  $\epsilon$ -amino groups) and the  $\alpha$ -amino group, all of which can link up with another ubiquitin molecule (B). Ubiquitin is also shown as a kite shape which is closer to the ubiquitin structure and shows conformations of ubiquitin within different chain isomers (C). Ubiquitin chains that are K48-linked adopt a more closed conformation (D) when compared with K63-linked and linear ubiquitin chains (E and F respectively).

The lysine involved in the poly-ubiquitin linkage and the number of ubiquitin molecules attached, are critical in determining the resulting role of the target protein. A chain of four or more, K48-linked poly-ubiquitin molecules will direct the target protein to the 26S proteasome for degradation<sup>82, 83</sup>. Alternatively, a K63-linked poly-ubiquitin chain may target the protein for involvement in DNA repair, signal transduction or endocytosis, for example by establishing new protein-protein interactions<sup>84</sup>. The biological significance of heteropolymeric ubiquitin chains has yet to be determined.

### **1.2.5 Deubiquitination of a target protein**

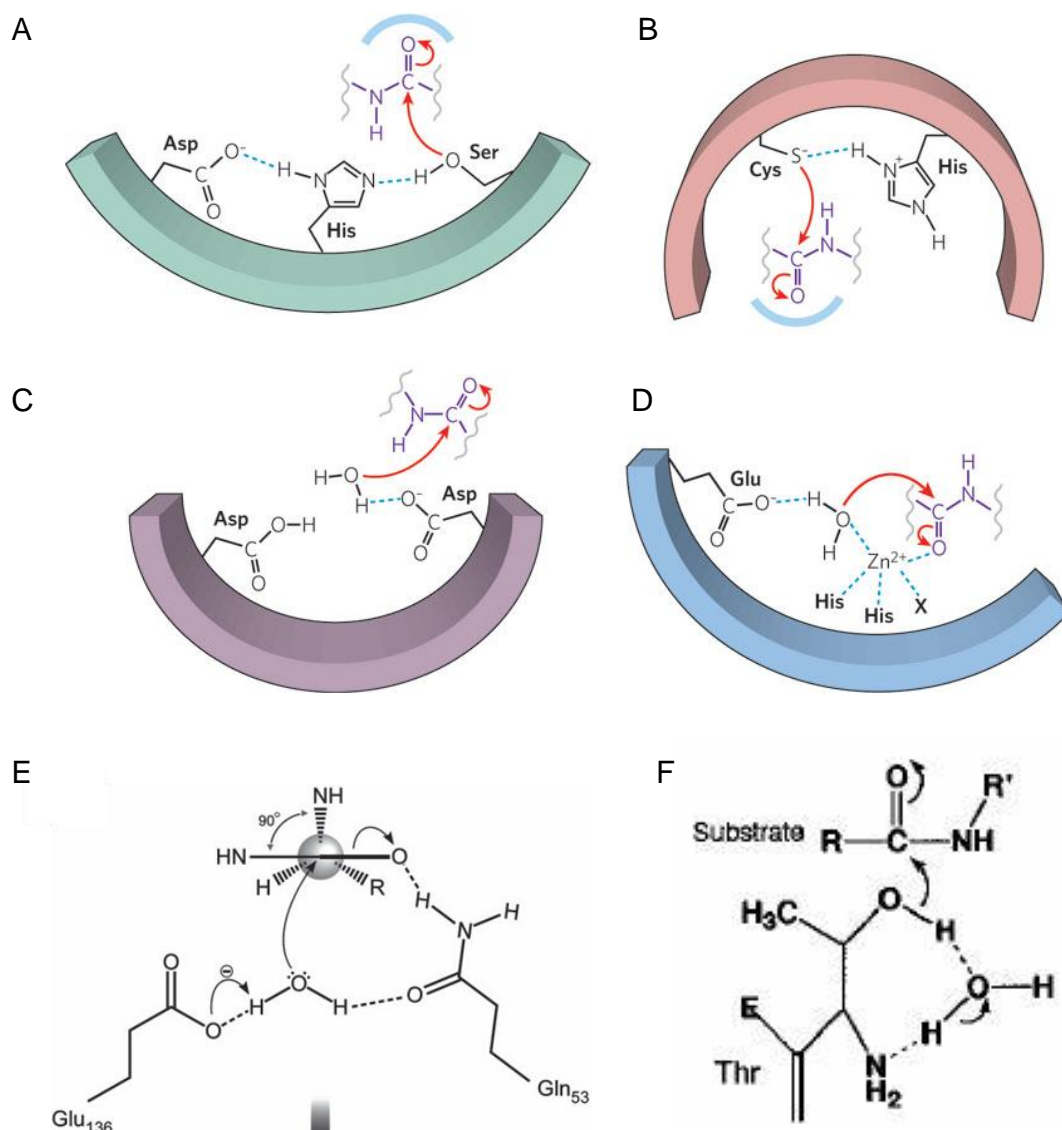
#### **1.2.5.1 The catalytic mechanisms of proteases**

Proteases are enzymes which cleave peptide bonds and are divided into six main groups: serine proteases, threonine proteases, cysteine proteases, aspartate proteases, metalloproteases and glutamic acid proteases. In order to cleave peptide bonds the protease uses up to three reactive functional groups within the active site, including a functional group with proton withdrawing, or deprotonating properties plus a nucleophilic group. If this process involves three (or two) amino acid residues then they are referred to as the catalytic triad (or diad).

These functional groups are in the side chains attached of the amino acids which form the DUB structure.

Proteases are grouped according to the residue which serves as a nucleophile or the residue that generates the nucleophile from a water molecule. The catalytic triad of serine proteases is made up of aspartate (sometimes asparagine), histidine and serine. Aspartate deprotonates a nitrogen within a five membered ring of histidine, which leads to the other nitrogen within the ring to deprotonate oxygen from the hydroxyl group of serine, which generates an oxygen nucleophile (Fig.1.7.A)<sup>85, 86</sup>. Cysteine proteases are not as well understood as serine proteases but are known to have histidine and cysteine as key amino acid residues within the active site, with sulfur from the cysteine residue forming the nucleophile (Fig.1.7.B). Cysteine and histidine have been shown to be able to function without asparagine<sup>87</sup>. The aspartate proteases use a catalytic diad made up of two aspartate amino acid residues. The oxygen from a hydroxyl group of aspartate is used to generate a nucleophile from a water molecule (Fig.1.7.C). Metalloproteases, zinc and an acid residue serve to make the oxygen of a water molecule nucleophilic (Fig.1.7.D). Reviewed by E. Erez *et al*<sup>17</sup>. In 2004 the glutamic acid protease was discovered in which the amino group of glutamine stabilises the substrate whilst the oxygen of the hydroxyl group of glutamate deprotonates a water molecule to generate

a nucleophile (Fig.1.7.E)<sup>20, 88</sup>. Later, the threonine proteases were discovered, interestingly these proteases use only an  $\alpha$ -amino group, threonine, to cleave proteins (Fig.1.7.F). Three  $\beta$ -subunits of the proteasome ( $\beta$ 1,  $\beta$ 2 and  $\beta$ 5) are known to have threonine catalytic sites<sup>21, 89</sup>. Threonine has two key functional groups, a free amino group to serve as a general base with a water molecule and a side chain hydroxyl group to form a nucleophile. The main focus for this study is on DUBs and ULPs, proteases that form part of the eukaryotic ubiquitin-mediated system, discussed in the next section.



**Fig.1.7** Cartoon to show the key amino acid residues within active sites of the six subgroups of proteases and how these serve to initiate nucleophilic attack on DUB substrates. An amino acid residue within the active site is made nucleophilic (A, B and F), which then carries out a nucleophilic attack on the carbonyl carbon of the substrate. A: The mechanism of serine DUBs, using a catalytic triad, in which aspartate (sometimes asparagine) uses its oxygen to deprotonate a neighbouring nitrogen within the ring of histidine. This leads to the remaining nitrogen within the ring of histidine deprotonating oxygen on serine leaving oxygen to serve as the nucleophile. B: A cysteine DUB is thought to operate under a similar mechanism to that of a serine DUB with sulphur serving as a nucleophile. Threonine proteases use two functional groups within threonine plus a water molecule to generate an oxygen nucleophile from the hydroxyl functional group of threonine. Aspartate proteases (C), metalloproteases (D) and glutamic acid proteases (E) involve processes in which a water molecule is made nucleophilic. Figures: 'A-D' from E. Erez *et al*<sup>17</sup>, 'E' from M. Fujinaga *et al*<sup>20</sup> and 'F' from Orłowski *et al*<sup>21</sup>.



#### 1.2.5.2 Tag removal as part of the Eukaryotic ubiquitin-mediated system

As discussed below, DUBs and ULPs are mostly cysteine proteases and some DUBs are metalloproteases. DUBs and ULPs operate with a similar catalytic mechanism to one another<sup>90</sup>. During deubiquitination, ubiquitin whilst covalently attached to the target protein, undergoes a nucleophilic attack by the DUB which eventually leads to a free ubiquitin molecule and target protein. DUBs and ULPs are modular, usually containing ubiquitin binding domains (UBD). There are more than twenty families of UBDs which adds to the complexity of the enzymes structure and therefore their binding to substrates and other proteins<sup>91</sup>.

Table 1.1 shows the seven classes of proteases encoded by eukaryotes and which remove protein tags in eukaryotes. Five of these classes are DUBs, the first of which is the ubiquitin C-terminal hydrolase (UCH), necessary to release ubiquitin from ubiquitin precursors<sup>92</sup>. The other four classes of DUBs are: ubiquitin specific proteases (USP) also known as ubiquitin processing proteases (UBP), DUBs carrying the ovarian tumour domain (OTU), metalloproteases (MM) and DUBs carrying the Josephin domain<sup>93-96</sup>. The Atg8 conjugation system is essential for forming autophagosomes in autophagy<sup>97</sup>. Atg4 is the DUB-like enzyme in this system. There are two groups of ULPs, the first is SUMO small ubiquitin-like modifier specific proteases (SENPs) and the second group of ULP is specific for either

NEDD8 or ISG15. The structures of some members of DUB and ULP groups have been solved, reviewed<sup>98</sup>.

Protease family	Classification of protease	Substrate		Example of DUBs/ULPs	
		Monomer type	Monomer / polymer	Eukaryote encoded	Pathogen encoded
C12	UCH	Ubiquitin	Ubiquitin precursor poly-ubiquitin	UCHL1, UCHL2 and UCHL3 deubiquitinate precursor ubiquitin <sup>99</sup> .	No eubacteria encoded deubiquitinating enzymes have been reported of this DUB class.
C19	USP/UBP	Ubiquitin	Mono-ubiquitin /K48-linked, K63-linked or K29-linked poly-ubiquitin	USPs are key DUBs in the UPS. CYLD inhibits the TNF-mediated NF- $\kappa$ B inflammatory response	
C54		Atg8	Atg8 monomer	Atg4 is essential for normal formation of an autophagosome <sup>100</sup> .	
C64, C85, C87 and C88	OTU	Ubiquitin	Mono-ubiquitin /K48-linked, K63-linked or K29-linked poly-ubiquitin	A20 and otubain1 OTU proteases both regulate the cell immune response <sup>101, 102</sup> .	
M67	MPN+/JAMM	Ubiquitin		A subunit of the proteasome, Rpn11/POH1 is a JAMM/MPN+ protease.	
C86	JD	Ubiquitin		Ataxin-3, a JD DUB was recently shown to edit ubiquitin chains of mixed linkage <sup>3</sup> . Ataxin-3 also associates with the proteasome and Rad23 which leads to the degradation of ubiquitinated proteins <sup>103, 104</sup> .	
C48	ULP	NEDD8 / ISG15	NEDD8 monomer / ISG15 monomer or polymer	SGN5 (CSN5) thought to be part of CSN complex which deNEDDylates cullins as a control over NF- $\kappa$ B signalling <sup>105</sup> .  UBP43 removes mono-ISG15 or poly-ISG15, a ubiquitin homologue. Cleavage maintains the integrity of the blood-brain barrier <sup>106</sup> .	Only 3 reported eubacterial DUBs, all of which inhibit the NF- $\kappa$ B inflammatory response:  1. SseL and AvrA are DUBs ( <i>Salmonella enterica</i> ) <sup>7, 8</sup> .  2. YopJ and YopP are DUBs and YopJ has also been shown to have deSUMOylating activity in plant hosts ( <i>Yersinia pestis</i> and <i>Yersinia enterocolitica</i> ) <sup>15, 107 9, 10</sup> .  3. ChlaDub1 and ChlaDub2 are DUBs which can also deNEDDylate ( <i>Chlamydia trachomatis</i> ) <sup>6, 108</sup> .
	ULP SENP 1-3	SUMO	Monomer or polymer	ULP1 and ULP2 Cleave SUMO-1 from I $\kappa$ B $\alpha$ and nuclear pore proteins (RanGAP1 and RANBP2) <sup>109</sup> .	

Table1.1

**Table.1.1** shows the 7 classes of proteases, with examples, for the removal of protein tags. The first 5 classes are DUBs: ubiquitin C-terminal hydrolase (UCH), ubiquitin specific proteases (USP), DUBs carrying the ovarian tumour domain (OTU), metalloproteases (MM), DUBs carrying the Josephin domain. The remaining 2 classes are ubiquitin-like proteases (ULPs), small ubiquitin-like modifier (SUMO) specific proteases (SENPs) and neural precursor cell expressed developmentally down-regulated protein 8 (NEDD8). Some examples of ULPs which have DUB activity and are encoded by eubacteria have been included.

There is evidence to suggest that DUBs are likely to be closely associated with E3 protein complexes responsible for ubiquitination. The human genome encodes ninety five DUBs, seventy five of which were investigated using a global proteomics analysis of DUBs and their complexes, by M. E. Sowa *et al*<sup>110</sup>. Twenty six of the DUBs in the study were found to be associated with proteins with a role in ubiquitination, including HECT E3 and Cullin-based E3 ligases. The close proximity between enzymes responsible for such opposing roles, reveals the tight controls over the targeting of proteins. Additionally, it was shown that six of the DUBs in the study were likely to interact with an AAA ATPase, VCP/p97, an enzyme involved in ubiquitin binding and the delivery of proteins to the proteasome.

It is clear that from the structures of some DUBs that certain domains can enable them to have additional roles to deubiquitinating. Not only can DUBs potentially deliver target proteins to the proteasome, one DUB, A20 has been shown to have E3 ligase activity. A20 uses its dual activities to edit K63-linked poly-ubiquitin chains to K48-linked poly-ubiquitin chains to negatively regulate the TNF-mediated NF- $\kappa$ B pathway<sup>111</sup>. When stimulated by TNF the TNFR forms protein complexes which lead to the ubiquitination and proteolysis of I $\kappa$ B $\alpha$ . This leaves NF- $\kappa$ B free to move to the nucleus and carry out transcription of genes associated with cell cycle arrest or apoptosis, as well as proteins

responsible for negative feed back on NF- $\kappa$ B signalling, I $\kappa$ B $\alpha$  and A20. A20 has an N-terminal OTU domain for DUB activity and seven zinc finger structures in the C-terminal domain for E3 ligase activity<sup>112, 113</sup>. A20 binds to a protein receptor-interacting protein (RIP), within the TNFR protein complex. First A20 deubiquitinates RIP removing the K63-linked poly-ubiquitin chain. A20 then ubiquitinates RIP this time with a K48-linked poly-ubiquitin chain, targeting RIP for disposal by the proteasome, resulting in the inhibition of the transcription of genes associated with cell cycle arrest or apoptosis<sup>114-116</sup>. This dual function of A20 is an indication of the complexity of the roles of DUBs within the eukaryotic ubiquitin-mediated system.

#### 1.2.5.3 Tag removal for virulence

As DUBs are often key enzymes within a signalling pathway, activating or deactivating key proteins within the pathway, this has lead to an interesting mode of virulence by pathogenic microbes. Over recent years there has been a steady increase in the number of reports of mimics of components of the eukaryotic ubiquitin-mediated system, including ULPs, encoded by eubacteria and viruses to take part in overriding the host cell systems for pathogenesis. However, viral proteins will not be considered here as the focus for this chapter is on eukaryotic ubiquitin-mediated systems and similar proteins of

eubacterial systems. There are six eubacteria encoded ULPs, AvrA, SseL, YopJ, YopP, ChlaDub1 and ChlaDub2 (Table 1.1). Despite being classed as ULPs they all primarily prefer ubiquitin as a substrate, although YopJ could possibly also act as a deSUMOylating enzyme and the ChlaDubs can deNEDDylate proteins. The subject of bacterial encoded DUBs for subverting host cell systems will be explored in Section 1.4 and revisited again in Chapter 6.

#### 1.2.5.4 Substrate specificity of DUBs or ULPs

DUBs and ULP levels in cells are controlled by transcriptional regulation, for example CYLD, the negative regulator of NF- $\kappa$ B signalling is also one of the products of this signalling pathway. With the varied topologies of ubiquitin substrates there are also DUBs and ULPs which are specific to certain types of these topologies. For example, the ULP, NEDP1 is specific for NEDD8 and as discussed previously (1.2.5.2) A20 is specific for K63-linked poly-ubiquitin<sup>117</sup>. The activity of DUBs and ULPs is also controlled in different ways.

Substrate-induced activity is a process which involves the binding of a substrate to an enzyme leading to allosteric changes taking place that realign key functional groups (discussed in 1.2.5.1) within the active site, to a conformation which favours catalysis. UCH-L3 is a DUB which

undergoes substrate-induced activity. The free structure for UCH-L3 was solved by S.C. Johnston and UCH-L3 bound to a suicide inhibitor by a group lead by H.L. Ploegh<sup>118, 119</sup>. It was noted that a loop crossed over the active site of UCH-L3 in the free form but when bound, this loop had formed an  $\alpha$ -helix and had moved to accommodate the substrate. H.L Ploegh's group then carried out an investigation to see if the loop was vital for catalytic activity or if it served to restrict the size of substrates, as it appears that the substrate has to enter through the loop to access the active site<sup>120</sup>. This study involved the development of a new technique which H.L. Ploegh's group named 'sortagging' which involved using an enzyme, sortase to partially digest points within the crossover loop which resulted in a loop that had slackened. After activation of this partially digested UCH-L3 with suicide inhibitors it was demonstrated that UCH-L3 maintained catalytic activity. In the same study, the loop was increased in length and it was found that the larger the loop the larger the substrate. These findings lead to the conclusion that the loop in UCH-L3 was mainly important for restriction on size of substrate.

The crossover loop is now thought to be a characteristic of all the UCH family of DUBs, however, there are also examples of members of other DUB or ULP families that have a crossover loop. For example, a member of the USP DUB family, CYLD is known to have a crossover



loop and it has been suggested that this may be important for determining specificity for K63-linked poly-ubiquitin chains as they adopt a more open conformation than K48-linked poly-ubiquitin<sup>121</sup>. The OTU domain DUB family has two members, Otu1 and otubaine-2, which are thought to have a crossover loop which forms a  $\beta$ -structure when active<sup>93</sup>. The ULP family also has an example of a protein, NEDP1 which has the crossover loop, only in this case it forms a  $\beta$ -structure in the active form<sup>117</sup>.

A final example of a DUB which undergoes substrate induced activity is the USP, isopeptidase T (IsoT). IsoT recycles mono-ubiquitin from unanchored poly-ubiquitin. IsoT has a zinc finger (ZNF) domain, two ubiquitin associated (UBA) domains and a USP domain. The ZNF domain has a pocket that is specific for the empty C-terminus of ubiquitin and only when this is bound can a conformational change occur that correctly aligns the catalytic triad to enable catalysis to take place<sup>122, 123</sup>.

DUBs have also been shown to become active when bound to a scaffolding protein or an adaptor protein. For example, the proteasome associated DUBs, Usp14, Uch37 and POH1 are only active when associated with the proteasome<sup>124, 125</sup>. DUBs have also been shown to become post-translationally modified which alters their active state. For

example, phosphorylation of A20 a USP which negatively modulates NF- $\kappa$ B signalling increases its activity, although it is not known if it is the ligase activity or the DUB activity is increased<sup>126</sup>. This contrasts with the phosphorylation of CYLD, another USP which negatively regulates NF- $\kappa$ B signalling as phosphorylation renders CYLD inactive<sup>127</sup>. Just as all proteins can be ubiquitinated for their disposal, so can DUBs and ULPs. However, the Josephine domain carrying DUB, ataxin-3 (ATN3) increases in activity when ubiquitinated<sup>128</sup>.

Ubiquitin and UBLs often operate in opposing directions in pathways in eukaryotes, therefore the integrity of some proteases must be ensured. For example, the SUMOylation of p53 by Mdm2 (1.2.2) leads to the activation of p53 but the ubiquitination of p53 leads to p53 being inactive. Therefore to maintain integrity of each opposing signalling pathway there is a need for tag recognition to be distinct for each pathway. This is achieved using the protease active site cleft in DUBs or ULPs, to recognise motifs in the last five to seven residues of the C-terminus of ubiquitin or the UBL respectively. The motif in ubiquitin is RLRGG, compared with NEDD8 which is RARLGG and the motif in SUMO is QQ/EQTGG<sup>109, 129</sup>. Many DUBs have also been shown to have ubiquitin binding domains to add to the specificity of the DUBs for ubiquitin, reviewed by D. Komander et al<sup>130</sup>. This will be discussed in more detail in the general discussion.

### **1.2.6 Disposal of a target protein**

Proteins targeted for disposal, from protein quality control systems or cell signalling are then delivered to a self-compartmentalised protease for proteolysis. Alternatively, proteins targeted as part of selective protein degradation via autophagy are transported in an autophagosome to undergo proteolysis in a lysosome<sup>100</sup>. Self-compartmentalised proteases are found in eukaryotic and prokaryotic cells. In eukaryotes these chambered proteases are in the cytosol, nucleus or are connected with the endoplasmic reticulum or the endoskeleton. In prokaryotes these proteases are thought to be cytosolic and associated with the cell membrane.

Typically, self-compartmentalised proteases have three main features: First, they are cylindrical; Second, they have a small pore restricting access to allow only unfolded proteins into the proteolytic chamber and third, Adenosine triphosphate (ATP) is hydrolysed by an ATPase-driven chaperone to unfold globular proteins and translocate the substrates into the proteolytic chamber. This produces peptide fragments of 10-15 amino acids. The structures of prokaryotic and eukaryotic self-compartmentalised proteases are compared in review articles<sup>131, 132</sup>.

In eukaryotes and archaea the 26S proteasome breaks down ubiquitin tagged proteins and is composed of: a 20S catalytic core, the ATPase

enzyme, 'ATPase associated with various cellular activities' (AAA) and a lid and base at each end, reviewed recently<sup>25</sup>. The 20S catalytic core is formed from  $\alpha$ -subunits which mainly form the structure of the proteasome and  $\beta$ -subunits which have active sites used in proteolysis. There are smaller self-compartmentalised proteases: ClpP, HslUV, FtsH and Lon found in prokaryotes and in the chloroplasts and mitochondria of eukaryotes<sup>131</sup>.

Proteases which are smaller than the eukaryotic 26S proteasome can form structures made up of one continuous protein as in the case of Lon or FtsH or composed of subunits. In *E. coli* these compartmentalised proteases are constructed of ClpP forming the proteolytic core and ClpX or ClpA as the ATPase or chaperone which work together to process SsrA-tagged proteins (Table 1.2)<sup>133</sup>. In these systems the proteins which attach the tags are SspB, RssB or the E3-like protein ClpS. ClpS has a putative type-2 binding site, normally found in the E3 ligase subgroup, N-recognins<sup>1, 2</sup>. *B. Subtilis* has ClpP as the catalytic core and ClpC as an ATPase or chaperone and the target proteins are tagged by MecA<sup>b</sup>.

Proteolytic core	ATPase	Protease family	Location in cell	Known tags or signals for proteolysis
ClpP	ClpX	Serine	Thought to be cytosolic.	SsrA bound to SspB
ClpP	ClpA	Serine		Polyphosphate tags or SsrA bound to SspB. ClpS (yljA) regulates proteolysis.
HSIV/ClpQ	HSIU/ClpY	Threonine		Recognition signals within the target protein.
Lon	-	Serine	Cytosol	Polyphosphate tags or proteins with exposed hydrophobic patches.
FtsH	-	Metallo-protease (zinc)	Anchored to inner membrane	Target proteins which are soluble and membrane associated and SsrA bound to SspB. Proteins targeted in response to heat shock / DNA damage.
ClpP Eubacteria, <i>B. subtilis</i>	ClpC	Serine	Cytosol	Recognition signals within the target protein.
20S Eubacteria, <i>M. tuberculosis</i>	Arc/Mpa	Threonine	Cytosol	Pup <sup>29</sup>
20S In archaea	PAN	Threonine	Cytosol	Unknown
20S In eukaryotes	19S	Threonine	Nucleus or cytosol, associated with ER.	Poly-ubiquitin

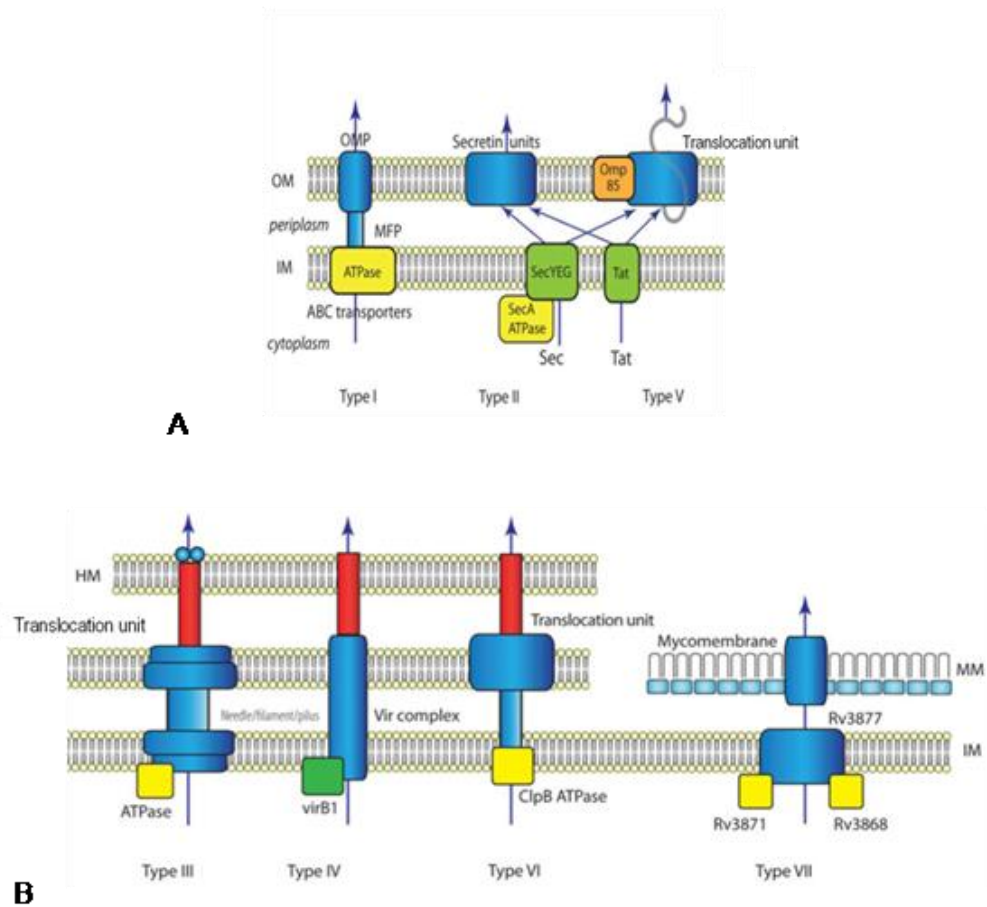
**Table.1.2** Self-compartmentalised proteases and associated tagging systems, found in eubacteria, Archaea and Eukaryotes<sup>1, 2</sup>. ClpP, HSIV/ClpQ, Lon and FtsH are all in *E. coli* and in the chloroplasts and mitochondria of eukaryotes.

### ***1.3 Microbes which subvert host cell systems***

In Section 1.2 the ways in which proteins are targeted to change their activity or for their disposal have been discussed as part of the eukaryotic ubiquitin-mediated system in comparison with what is currently understood about similar systems in prokaryotes. The processes discussed in section 1.2 were those used as part of a signalling or protein quality control system important for cell survival. In this section it would be pertinent to change focus now to the proteins encoded by eubacteria as either virulence factors to subvert host cell systems, or to form a symbiotic relationship with the host.

Virulence factors are proteins which are used by pathogenic eubacteria to over ride host cell systems. These virulence factors are released using needle-like structures so that the host cells can receive them. Virulence factors and the needle-like structures are coded for in genetic islands, which are either mobile in plasmids or as distinct islands on the chromosome. Although in a few cases bacteria have been reported to have these secretion systems, they are normally only found in pathogenic bacteria, in which the genomic island is referred to as a pathogenicity island. Only pathogenic *E. coli* have been reported to have these secretion systems.

There are seven main types of needle like structures called Type I – Type VII secretory systems. Fig.1.7 shows these secretory systems, taken from a recent review by T. Tseng *et al*<sup>12</sup>. The Type II secretory system differs from Type I and Type III secretory systems as it modifies proteins within the periplasm on route. The pathogenic eubacteria which encode DUBs as virulence factors use the Type III secretory system, which has been shown to inject the host directly with the effector proteins<sup>134-136</sup>.



**Fig.1.7** A schematic diagram of Type I, Type II, Type IV (A), Type III and Type, IV-VII (B) secretory systems in eubacteria taken from a recent review by T. Tseng *et al*<sup>12</sup>. Hydrolysis of ATP is used to secrete proteins across the membranes. Type I and Type III systems are similar as they involve the release of proteins in one step across the periplasm. HM: Host membrane; OM: outer membrane; IM: inner membrane; MM: mycomembrane; MFP: membrane fusion protein. ATPases and chaperones are shown in yellow.

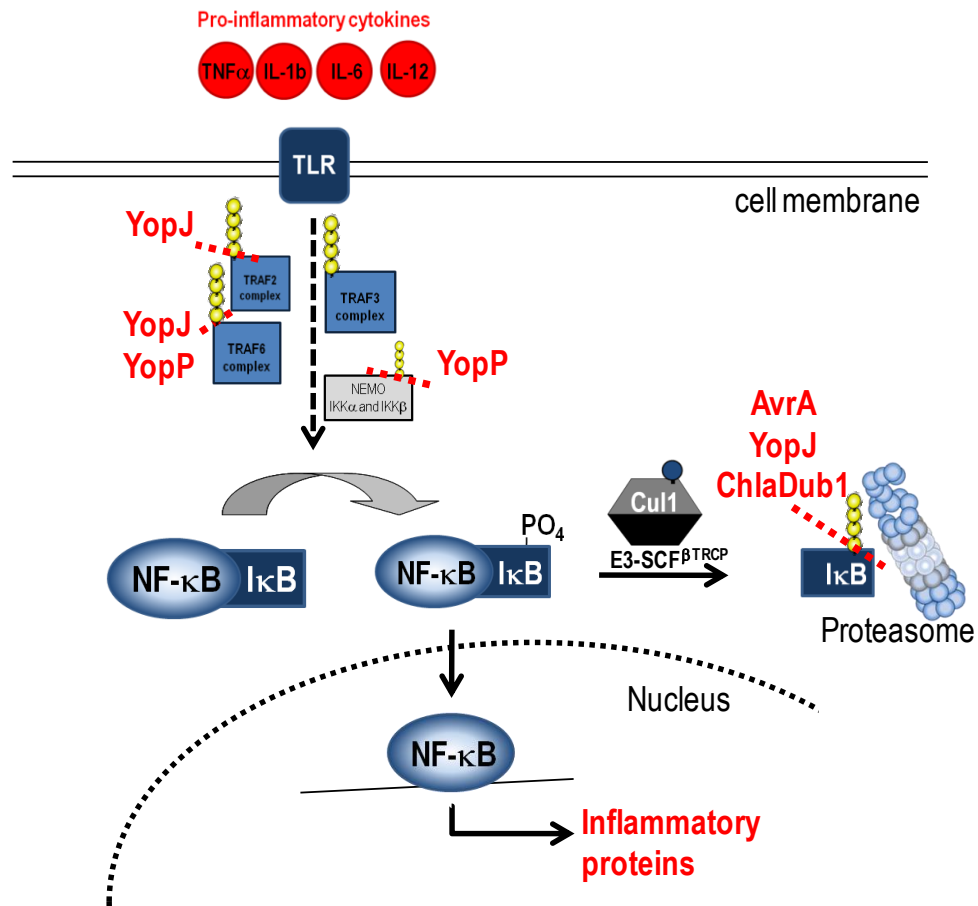


Eubacteria have been shown to express mimics of the eukaryotic ubiquitin-mediated system which modulate the host ubiquitin-mediated processes as part of virulence. For example *P. syringae pathover* infects plants, using an E3 mimic, AvrPtoB, as a virulence factor to disable the host cell's self-destruct signal to enter into an apoptotic state, therefore leaving the bacterium to replicate undisturbed<sup>137</sup>. As mentioned previously (1.2.5) some pathogenic eubacteria have been shown to express DUBs (with E3 mimics in some cases) which are also used for virulence.

To date, there have been only 3 pathogenic eubacteria reported to express DUBs and do so as part of a Type III secretory system coded for on a pathogenicity island. These DUBs serve as virulence factors, inhibiting the activity of NF- $\kappa$ B in the eukaryotic interleukin-mediated NF- $\kappa$ B signalling pathway. The DUBs achieve this by deconjugating K63-linked or K48-linked poly-ubiquitin chains from key proteins activators within the NF- $\kappa$ B signalling pathway (Fig.1.2). This results in the inhibition of the inflammatory response.

SseL and AvrA are DUBs encoded by *Salmonella enterica*<sup>7, 8</sup>. In some cases the target proteins for this DUB activity by the virulence factors have been established, shown in Fig.1.8. YopJ and YopP are DUBs expressed in *Yersinia pestis* and *Yersinia enterocolitica* respectively<sup>9, 15</sup>.

<sup>107</sup>. YopJ and AvrA target I $\kappa$ B, removing a K48-linked poly-ubiquitin chain so rescuing I $\kappa$ B from disposal by the proteasome. YopJ also targets TRAF2 and TRAF6, removing K63-linked poly-ubiquitin chains which prevents the phosphorylation of I $\kappa$ B and consequently maintains the NF- $\kappa$ B in an inactive state, preventing expression of inflammatory proteins<sup>15</sup>. YopP achieves the same result as YopJ but by removing K63-linked poly-ubiquitin chains from TRAF6 and NEMO including two other proteins in complex with NEMO, IKK $\alpha$  and IKK $\beta$ <sup>9</sup>. ChlaDub1 and ChlaDub2 are DUBs which can also deNEDDylate and are expressed in *Chlamydia trachomatis*<sup>6, 108</sup>. ChlaDub1 has been shown to target I $\kappa$ B removing a K48-linked poly-ubiquitin chain<sup>6</sup>.



**Fig.1.8** A simplified schematic view of the interleukin-mediated NF- $\kappa$ B signalling pathway to show key protein complexes targeted by pathogenic eubacteria using DUBs as virulence factors. The NF- $\kappa$ B signalling pathway is activated by pro-inflammatory cytokines released as part of the immune response. DUBs are shown in red and red dashed lines show deubiquitinating activity. AvrA, YopJ and YopP are DUBs expressed by *Salmonella enterica*, *Yersinia pseudotuberculosis* and *Yersinia colitica*, respectively<sup>15</sup>. The result of this inhibition of the NF- $\kappa$ B pathway is to inhibit the expression of inflammatory proteins which would normally attack bacteria as part of the immune response.

#### **1.4 DUB activity observed in non pathogenic *E. coli***

In the previous section the expression of mimics of the eukaryotic ubiquitin-mediated system, as virulence factors in eubacteria was discussed. In the current section the occurrence of a possible DUB activity in non-pathogenic *E. coli* will be considered. There have been a few observations in the literature of a possible DUB activity in *E. coli* as well as the discovery of a ULP with DUB activity in *E. coli*. It is therefore interesting to consider what functional role a DUB may have in a bacterium that lacks the secretory systems used by pathogenic eubacteria for virulence.

There have been a few reports observing intrinsic, low level DUB activity in the eubacterium, *E. coli*. One group reported DUB activity in *E. coli* on two occasions, once when using bacteria to express penta-ubiquitin, and immunoblotted for ubiquitin, which revealed the break down products from penta-ubiquitin: mono-, di-, tri- and tetra-ubiquitin<sup>3</sup>. The second time, a possible DUB activity was observed again, while developing a robust ubiquitin fusion protein expression system in *E. coli*<sup>138</sup>. DUB activity was also observed by A. M. Catanzariti and colleagues and demonstrated to be the result of a specific cleavage by sequencing<sup>4</sup>. Another group, A. Ciechanover and colleagues reported the eukaryotic elongation factor, EF-1 $\alpha$ , a protein identical to their protein of interest, factor hedva (FH)<sup>139</sup>. The authors also mentioned a

DUB activity associated with a prokaryotic elongation factor, EF-Tu, a protein which co-migrated with the activity. The structure and function of elongation factors in the delivery of charged tRNA to ribosomes during translation has been fully described<sup>140-142</sup>.

Whilst our study was in progress, A. Catic reported the discovery of a ULP, *elaD* expressed in non-pathogenic *E. coli*, present in *E. coli* substrain K12 and intestinal pathogenic strains but absent from extraintestinal strains<sup>14</sup>. Bioinformatics analysis was used to search for new members of the clan containing ubiquitin and ubiquitin-like proteases (clan CE) in viruses, bacteria and eukaryotes, which revealed the protein, *elaD* as a potential protease. Two biochemical tests were then carried out *in vitro*, which showed that GST-*elaD* displayed protease activity associated with a DUB or ULP. The first biochemical test involved reacting GST-*elaD* with substrates termed suicide inhibitors, which were types of Michael acceptors, vinyl methyl ester (VME) or vinyl sulphone (VS). These Michael acceptors were attached C-terminal to: ubiquitin, Nedd8 or SUMO. GST-*elaD* formed an adduct with UbVME and to a small extent with Nedd8VS but not SUMOVS. The second biochemical test involved incubating GST-*elaD* with a fluorogenic substrate. GST-*elaD* deconjugated Ub-AMC but not SUMO1-AMC or Nedd8-AMC.

The discovery of a ULP in *E. coli*, together with the observations by three other groups, of a DUB activity in lab strains of *E. coli* lead one to speculate briefly upon the possible function for a DUB activity in non pathogenic strains of *E. coli*. Eubacteria may possess a ubiquitin-like system. However, as mentioned previously (1.2.1), eubacteria do not express ubiquitin and the only eubacterium so far, to be shown to express a ubiquitin homologue, Pup (6.9 kDa) is *Mycobacterium tuberculosis*<sup>29</sup>. It is also unlikely that a dePupylating enzyme would use the same mechanism as a DUB as Pup is attached to the targeting protein differently compared with ubiquitin. In *E. coli* however, it is still a possibility that a UBL and more ULPs have yet to be discovered. Alternatively, it is possible that as *E. coli* is a commensal bacterium found in the large intestine that the function of the DUB activity is associated with modifying the immune response in some way, to enable it to survive. As commensal (non-pathogenic) *E. coli* do not have genetic islands coding for secretion systems, this comes with the caveat that *elaD* must be a secreted effector and needs to enter the host cell in some way.

### ***1.5 The main aims for this study***

The work in this study started in response to the observation of a ubiquitin-fusion processing activity associated with an undocumented

DE3 substrain of *E. coli*, by other members of this group (H. Jefferey and I. Kerr) see section 3.2 for further details. This discovery and the observations made in the literature indicated that there may be an intrinsic DUB-like activity in *E. coli* or in some of its substrains. At that time, no DUB or ULP had been reported in *E. coli*. These findings initiated the current study, to characterise the ubiquitin-fusion processing activity in *E. coli*.

There were four main aims of this study. The first two aims were to establish that the ubiquitin-fusion processing activity previously observed by H. Jefferey and I. Kerr was demonstrated with the *E. coli* substrain Rosetta™2(DE3), then to determine the nature of the cleavage as specific DUB-like or non-specific. The third aim of this study was, once the type of cleavage was known, to investigate the specificity of the ubiquitin-fusion processing activity expressed in *E. coli* substrain Rosetta™2(DE3). The final aim was to take a candidate approach to identify the enzyme or enzymes responsible for the DUB activity observed in *E. coli* substrain Rosetta™2(DE3).

## **2 Materials and methods**

### ***2.1 Cells, media and preparation of lysates***

#### **2.1.1 Cells and media**

*E. coli* substrains used in this study were either XL10 Gold or DH5 $\alpha$  (to generate clones) and Rosetta™2(DE3) to express proteins. The strains were grown (37°C) with agitation in Luria-Bertani (LB) broth (1% tryptone, 0.5 % yeast extract, 1 % NaCl) (SIGMA-Aldrich) to an optical density = 0.6 – 0.8. Where appropriate ampicillin (100  $\mu$ g/ml) was used to select for antibiotic resistance. IPTG (50  $\mu$ M – 200  $\mu$ M) was used to induce protein expression. *E. coli* strain, K12, MG1655, substrain, JIG182 was used initially to generate the *elaD* knockout and was grown (30°C) with agitation in LB broth to an optical density = 0.6-0.8. Where appropriate ampicillin (100  $\mu$ g/ml) was used to select for antibiotic resistance. Apramycin (35 $\mu$ g/ml) was used to select for the *elaD* knockout colony. Protein expression was induced by 0.2 % (w/v) arabinose.

*Lactococcus lactis* strain, NZ9000 was used to express the (His)<sub>6</sub>-Ub-LmrC construct (made by J. Dorrian) using pNZ8048 which carries a nisin promoter to control protein expression. Nisin is a protein which is also expressed in *L. lactis* (see below for the generation of nisin). *L.*



*lactis* was grown (30°C) without agitation in M17 medium (Merck, Darmstadt, Germany) supplemented with 0.5% (w/v) glucose to an optical density = 0.6. Chloramphenicol (5 µg/ml) was used to select for antibiotic resistance. Protein expression was induced by addition of a culture supernatant from the nisin-expressing strain, nz9700 at a volumetric ratio of spent media : culture to be induced, 1: 8000.

Enhanced nisin expression was obtained using plasmid, nz9700 in *L. lactis*. Cultures were grown as described above with no chloramphenicol until optical density = 0.9. A clarified lysate was prepared as described below and the resulting spent media was used for protein induction. Due to variations in concentration of nisin between clarified lysates it was necessary to find the volume of nisin to volume of culture required for optimum protein induction, which in this study was 1/8000.

### **2.1.2 Preparation of cell lysates**

Cell cultures were pelleted, resuspended in TBS (150 mM NaCl, 10 mM Tris) unless indicated otherwise and sonicated for 6 x 10s bursts (4°C) followed by centrifugation. The resulting supernatant was then used for the purposes described later.

## ***2.2 Oligonucleotides used in the study and standard nucleic acid techniques***

### **2.2.1 Oligonucleotides used in this study**

See Table 2.1.

### **2.2.2 Plasmid preparation**

DNA was extracted from the cell lysate supernatant, using plasmid preparation kit: QIAprepSpin (QIAGEN).

### **2.2.3 Obtaining genomic DNA to use as a DNA template in PCR**

A sterile pipette tip was used to scrape a glycerol stock of the bacterial culture, in to 50 µl sterile distilled water. The resulting cell solution was then mixed and left at 4°C for 20 min.

### **2.2.4 Polymerase chain reaction (PCR) and site directed mutagenesis (SDM)**

#### **2.2.4.1 PCR using Pfu DNA polymerase (Promega)**

PCRs were carried out using Pfu buffer (20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 0.1 % (v/v) Triton X-100 and 0.1 mg/ml nuclease-free BSA), the reactants shown in Table 2.2. The PCRs were carried out in a thermal cycler using cycling conditions described in Table 2.3.

#### 2.2.4.2 PCR using GoTaq<sup>®</sup> DNA Polymerase (Promega)

PCRs were carried out using 5X Colourless GoTaq<sup>®</sup> Reaction Buffer (Part# M792A, M792B): Proprietary formulation supplied at pH 8.5. This buffer contains 7.5 mM magnesium. Buffer formula not specified. The PCRs were carried out in a thermal cycler using the reactants in Table 2.2 and conditions described in Table 2.3.

#### 2.2.4.3 SDM

##### 2.2.4.3.1 SDM using *Pfu* DNA polymerase (Promega):

PCR based SDM using the methods described in the QuickChange Instruction Manual, Revision A.01 (Stratagene), *Pfu* DNA polymerase (2.2.4.3.1) and the cycling parameters described in Table 2.4.

#### 2.2.5 Further treatment of amplified DNA products:

##### 2.2.5.1 PCR products for cloning:

PCR products were either purified using QIAquick PCR Purification Kit (QIAGEN) or gel purified by electrophoresing with DNA standards 200–10000 bp, SmartLadder (Eurogentec) on a 1% (w/v) agarose gel, 70 V. Resulting bands were visualised using UV light and excised using sterile a razor blade and the DNA was isolated using QIAquick Gel Extraction kit (QIAGEN).

#### 2.2.5.2 PCR products for SDM or to generate *elaD* knock-outs:

PCR products were incubated with Dpn1 (10 u), 37°C for 1 hr to digest the parental DNA template and either introduced in to bacterial cells or was ethanol precipitated for creating *elaD* knock-outs (2.2.8).

#### 2.2.6 Screening for colonies using colony PCR

Use of colony PCR to screen *E. coli*. Substrains, K12 and Rosetta<sup>TM</sup>2(DE3) for the *elaD* knock-out. Colonies were purified and used to set up LB 1.5% (w/v) Agar Select either without antibiotic, with ampicillin (100 µg/ml) or with apramycin (35 µg/ml) as follows. A sterile pipette tip was used to touch a purified colony and then to inoculate the three types of LB Agar select plates above as well as to set up a template for colony PCR. The Ampicillin plates were used to check for successful curing of the pKD46 plasmid and the non selecting plates were used check cell viability in case there were no colonies on the other two types of plates. PCR was carried out using GoTaq<sup>®</sup> DNA polymerase, primers, *ElaDCHF*, *ElaDCHR* and method described in 2.2.4.2.

#### 2.2.7 Restriction endonuclease digestion

Purified DNA (400 ng) was digested with the appropriate pair of restriction enzymes (5 u of each), 1-2 hr, 37°C and run with DNA

standards 200–10000 bp, SmartLadder (Eurogentec) on an 1% (w/v) agarose gel, 70 V.

#### **2.2.8 DNA precipitation for use in generating *elaD* knock-out in K12**

Chilled ethanol was added to a mixture of sodium acetate (3 M pH 5.2) and DNA to volume ratio, 27.5:1:10, mixed then incubated (-20°C) 25 min. The DNA solution was centrifuged (14,000g), 4°C for 30 min and the resulting pellet washed in 70% ethanol (chilled), air dried over night and resuspended in sterile water.

#### **2.2.9 Ligation**

Gel purified double digested inserts (I) and the appropriately digested vector (v) were usually ligated in volume ratios, I : v; 9:3, 6:3, 3:3, 3:9; The DNA was ligated in a 20 µl reaction mix containing: ligase (1.5 u), 30 mM Tris-HCl (pH 7.8), 10 mM MgCl<sub>2</sub>, 10 mM DTT and 1 mM ATP, 2 hr at 20°C

#### **2.2.10 Phage 1 (P1) transduction**

The cell pellet from a 4 ml culture of the recipient strain, Rosetta™2(DE3) was resuspended in 1 ml of MC Buffer (100 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 mM CaCl<sub>2</sub>·2H<sub>2</sub>O). The recipient cell solution (200 µl) was incubated with Phage1 lysate, 37°C for 30 min and the Phage1

infection of the recipient strain was stopped using 200  $\mu$ l sodium citrate (1 M). The infected recipient culture was then mixed with an overlay agar (1 % (w/v) Bactotryptone, 0.5 % (w/v) yeast extract, 171 mM NaCl, 2 mM NaOH) and poured over a 1.5 % (w/v) LB Agar select plate containing apramycin (35  $\mu$ g/ml).

### **2.2.11 DNA electrophoresis**

DNA was electrophoresed with DNA standards 200–10000 bp, SmartLadder (Eurogentec) on a 1% (w/v) agarose gel, 70 V.

## **2.3 Transformation**

### **2.3.1 Transformation of chemically competent bacteria**

#### **2.3.1.1 Preparation of chemically competent bacteria:**

Bacterial cells were grown (30°C) with agitation in LB broth to optical density = 0.5-0.8. The cells were pelleted down (4°C) by centrifugation (14,000 g) for 5 min, washed in 0.1 M  $\text{MgCl}_2$  and incubated in 0.1 M  $\text{CaCl}_2$  and incubated (4°C) for 20 min. After a further centrifugation (14,000 g) for 5 min the cells were resuspended in chilled 16% (v/v) glycerol in 0.1 M  $\text{CaCl}_2$ .

#### **2.3.1.2. Transformation of competent bacteria**

Where appropriate, competent DH5 $\alpha$  and Rosetta™2(DE3) cells (100  $\mu$ l) were incubated with plasmids (10  $\mu$ l), Dpn1 digested SDM

product (5 µl) or ligated vector (10 µl), as follows; 30 min (4°C); 30 s (42°C); 2 min (4°C). LB was added then incubated for 1hr with agitation (37°C). Controls included LB only to check for media infections and cells (- DNA) to check for cell infections. Cells were grown (37°C) on 1.5 % (w/v) LB Agar Select.

### **2.3.2 Transformation of bacteria using electroporation in generating the *elaD* gene knock-out JIG182Δ*elaD*<sup>Apr</sup> used ultimately to make Rosetta™ 2(DE3)Δ*elaD*<sup>Apr</sup>**

The *elaD* gene knock-out was first created in *E. coli* substrain K12, MG1655 (JIG182) using homologous recombination to replace *elaD* with an apramycin resistance cassette (*Apr*<sup>R</sup>). This *Apr*<sup>R</sup> cassette (replacing *elaD*) was then transferred to Rosetta™ 2(DE3) using Phage1 transduction.

#### **2.3.2.1 Preparation of bacteria for electroporation to generate JIG182Δ*elaD*<sup>Apr</sup>:**

*E. coli* substrain K12, MG1655 (JIG182) was grown in a 40 ml culture (30°C) with agitation in SOB, ampicillin (100 µg/ml) to optical density = 0.6. JIG182 were then washed four times in 1 mM HEPES and resuspended in 100 µl 1 mM HEPES.

#### **2.3.2.2 Electroporation of K12 to make JIG182Δ*elaD*<sup>Apr</sup>:**

Purified DNA (~2 µg ) was added to 67 µl of washed bacterial cells in a 0.1 cm cuvette and the BIORAD Gene Pulser Xcell electroporation system supplied a V pulse, 1.7 KV. SOC was added to the cells immediately which were then grown on 1.5 % (w/v) LB Agar Select, with apramycin (35 µg/ml), at 42°C to cure the pKD46 plasmid.

#### 2.3.2.3 Induction of λ Red proteins involved in generating JIG182ΔelaD<sup>Apr</sup>:

*E. coli* substrain K12, MG1655 (JIG182), carrying pKD46 which codes for the λ Red enzymes responsible for carrying out homologous recombination were used to generate JIG182ΔelaD<sup>Apr</sup>. JIG182 was grown (30°C) in a 40 ml culture, including 0.2% (w/v) arabinose, to induce protein expression of the enzymes in the λ Red system, with agitation in SOB, ampicillin (100 µg/ml) to optical density = 0.6.

## **2.4 Standard protein techniques**

### **2.4.1 Affinity purification of (His)<sub>6</sub>-tagged recombinant proteins**

A bacterial culture, typically 1 L, was grown up as described, 2.1.1 and lysates prepared (10 ml) detailed in 2.1.2, but using Wash and Equilibration (WE) buffer (50 mM sodium phosphate, 300 mM sodium chloride and 10 mM imidazole ACS reagent) in place of TBS (150 mM NaCl, 10 mM Tris). Protein purification by either Ni<sup>2+</sup> or Co<sup>2+</sup> affinity



chromatography: 10 ml lysate was then incubated (4°C) on a rotator with 1 ml slurry of either Ni<sup>2+</sup> or Co<sup>2+</sup> affinity resin (Sigma) for 30 min – 60 min. The specifically bound, immobilised proteins were washed in WE Buffer, then eluted in Elution buffer (50 mM sodium phosphate, 300 mM sodium chloride and 250 mM imidazole).

## **2.4.2 Affinity purification of GST-tagged recombinant proteins**

### **2.4.2.1 Capture of GST-tagged proteins on Glutathione-Sepharose<sup>TM</sup>4B beads**

A bacterial culture, typically 1 L, was grown up as described, 2.1.1 and lysates prepared detailed in 2.1.2. Lysate (10 ml) was then incubated (4°C) on a rotator with a slurry (1 ml) of Glutathione-Sepharose<sup>TM</sup>4B beads (Amersham) for 1 hr. The specifically bound, immobilised proteins were washed in TBS supplemented with Triton X-100 (TBST) (150 mM NaCl, 10 mM Tris 0.1% (v/v) Triton X-100 pH 7.5). Alternative buffer washes to TBST were also used in an attempt to wash off GroEL (Chapter 4): High salt (2 M NaCl, 50 mM Tris 0.1% (v/v) Triton X-100 pH 7.5), B-PER (100%) and ATP buffer (2 mM ATP, 10 mM MgSO<sub>4</sub>, 50 mM Tris-HCl pH 7.4). These buffers were not used for DUB assays as they were not appropriate for protein-protein interaction studies.

#### 2.4.2.2 Eluting GST-tagged proteins from Glutathione-Sepharose<sup>TM</sup>4B beads

The specifically bound, immobilised proteins on Glutathione-Sepharose<sup>TM</sup>4B beads (captured as described in 2.4.2.2 using TBST washes) were incubated (4°C) 30 min rotating in 20 mM reduced glutathione in TBS (pH 7.5) (1 ml).

#### 2.4.2.3 Thrombin cleavage to remove the GST-tag from GST-tagged proteins

The specifically bound, immobilised proteins on Glutathione-Sepharose<sup>TM</sup>4B beads (captured as described in 2.4.2.2 using TBST washes) were washed in 10 ml thrombin cleavage buffer (TCB) (19.8 mM Tris base, 149 mM NaCl and 2.5 mM CaCl<sub>2</sub>·2H<sub>2</sub>O (pH 8.4)) then incubated (4°C) 30 min rotating in 1 ml thrombin (5u/500 µl dry beads) in TCB.

### **2.4.3 Dialysis of purified proteins**

Purified proteins were incubated in dialysis tubing (4°C), stirring over night in TBS with 0.1 mM DTT (pH 7.5) (2 L).

### **2.4.4 Generation of ZNF-Sepharose 4B beads**

Cyanogen bromide activated Sepharose 4B beads were incubated with 1 mM HCl solution for 15 minutes then washed with coupling buffer (100mM NaHCO<sub>3</sub>, 500mM NaCl, pH8.3 (HCl)). Thrombin cleaved

ZNF216 (wild type) was attached to the Sepharose beads by incubating rotating (4°C) for 2-3 hours then washed in 1M ethanolamine (pH8) and incubated (4°C) in 1 M ethanolamine overnight. This was then followed by 3 washes in coupling buffer, one wash with acetate buffer (100mM CH<sub>3</sub>COONa, 500mM NaCl, pH4), followed by three washes with Tris buffer (100mM Tris (hydroxymethyl) methylamine, 500mM NaCl, pH8), resulting in approximately 8-12mg protein/ml of Sepharose.

#### **2.4.4 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)**

A gradient gel, 5-20% (w/v) acrylamide was made up using a gradient mixture; 5% (v/v) acrylamide (0.25 M Tris, and 0.1% (v/v) SDS); 20% (v/v) acrylamide (0.25 M Tris, 1.92 M glycine and 0.1% (v/v) SDS). Protein samples were incubated (95°C) for 7 min in gel loading buffer (GLB) (0.15 M Tris, 8 M urea, 2.5% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, 3% (w/v) DTT, 0.1% (w/v) bromophenol blue pH 6.8) and run in electrode buffer (25 mM Tris, 186 mM glycine and 0.1% (w/v) SDS) at 40 mA.

#### **2.4.5 Coomassie staining and destaining**

Staining and destaining gels: Gels were stained by incubating (1 hr) in Coomassie Blue stain (50% (v/v) methanol, 20% (v/v) glacial acetic acid

and 0.12% (w/v) Coomassie brilliant blue R-250) and destained by incubating (over night) in Destain solution (10% (v/v) methanol and 10% (v/v) glacial acetic acid).

#### **2.4.6 Transfer of proteins to nitrocellulose membrane for N-terminal sequencing and for Western blotting (WB)**

Proteins were transferred on to nitrocellulose membrane (NC) (Amersham), in transfer buffer (0.025 M Tris, 0.192 M glycine and 20% (v/v) methanol pH 8.3), at 40 mA over night, unless stated otherwise. The specifically labelled proteins were visualised by incubating the membrane (2 min) with Western Lightning® Plus-ECL, Enhanced Chemiluminescence Substrate (PerkinElmer), and was used to expose photographic film.

##### **2.4.6.1 Preparation of proteins for N-terminal sequencing (Edman degradation)**

Proteins were electrophoresed as described in (2.4.3) then the gel was incubated (15 min), rocking in transfer buffer (CAPs (10 mM) methanol (20%) (v/v) pH 11.0), to clear excess glycine. Proteins were then transferred on to Polyvinylidene fluoride (PVDF) from Amersham, in transfer buffer (CAPs (10 mM) methanol (20%) (v/v) pH 11.0) at 40 mA over night. The resulting membrane was washed in distilled water,

incubated (2 min) in Amido Black solution (16 mM Amido black, 7.5% (v/v), glacial acetic acid 20% (v/v) methanol), shaking, to reveal the proteins, then washed in distilled water. The resulting blot was then submitted for analysis (Protein and Nucleic Acid Chemistry Facility, University of Cambridge).

#### 2.4.6.2 WB for (His)<sub>6</sub>-tagged recombinant proteins

The NC membrane was washed in TBST (25 mM TBS, Tween (1:1000 v/v)) and blocked in 5% (w/v) Marvel in 0.25% (v/v) TBST. The NC membrane was washed in TBST before and after incubation with HRP-labelled mouse anti-hexahistidine antibody (R&D Systems) in TBST (1:5000) for 1 hr. The HRP label on the primary antibody negated the secondary antibody step normally used in Western blotting.

#### 2.4.6.3 WB for ubiquitin

The NC membrane was autoclaved (20 min) and blocked (1 hr) with 5% (w/v) Marvel in TBS before incubating in anti-ubiquitin primary polyclonal antibody from a rabbit, in 5% (w/v) Marvel (1:1000). After washing in TBS (10 mM Tris and 150 mM sodium chloride, pH 7.5) the NC membrane was incubated in anti-rabbit polyclonal antibody from swine in 5% (w/v) Marvel (1:2000), for 1 hr, then washed in TBS.

#### 2.4.6.4 WB for GST-tagged recombinant proteins

The NC membrane was washed in TBST (25 mM TBS, Tween (1:1000 v/v)) and blocked in 5% (w/v) Marvel in TBST. The NC membrane was washed in TBST after incubation with the HRP labelled mouse anti-GST antibody (Universal Biologicals) (1:10 000) in 5% (w/v) Marvel in TBST for 1 hr. The HRP label on the primary antibody negated the secondary antibody step normally used in Western blotting.

#### 2.4.6.5 WB for SUMO1

The NC membrane was washed in TBS and blocked in 5% (w/v) Marvel in TBS with 0.1% (v/v) Tween. The NC membrane was washed in TBS after incubation with the SUMO1 antibody (mouse anti-GMP-1, Zymed Laboratories) in 5% (w/v) Marvel in TBS with 0.1% (v/v) Tween (1:1000), 4°C, 15 hr.

### ***2.5 Deubiquitination assays***

#### **(His)<sub>6</sub>-Ub-LmrC deconjugation assay**

*E. coli* substrain Rosetta<sup>TM</sup>2(DE3) cell lysate, unpurified proteins: GST-elaD, GST, PGP9.5 or purified proteins: USP2-core M231 were incubated (37°C) with either purified (2.4.1) or unpurified (2.1.2) (His)<sub>6</sub>-Ub-LmrC protein (generated in *Lactococcus lactis*) in TBS with 0.5 mM

DTT (pH 7.5) for approximately 15 hr. PGP9.5 and M231 were included as positive controls for DUB activity. The reaction was stopped using gel loading buffer and the proteins were analysed by western blotting for ubiquitin.

### **Poly-ubiquitin deconjugation assays**

GST-tagged proteins were purified using glutathione-Sepharose 4B beads and while still attached to beads were then incubated (37°C) with poly-ubiquitin substrates in TBS with 0.5 mM DTT (pH 7.5) for approximately 15 hr. Either USP2-core or M231 were included as a DUB to serve as a positive control for DUB activity. The reaction was stopped using gel loading buffer and the proteins were analysed by western blotting for ubiquitin.

## ***2.6 Cloning methods and generating mutants***

### **2.6.1 Cloning recombinant (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-SUMO1-LmrC or (His)<sub>6</sub>-Ub-EGFP using a modified pHUE vector**

DNAs encoding human NEDD8 (GI:4738) and human SUMO1 (GI:4507801) were supplied in a vector (pZHFN, Invitrogen) by S. Dawson; the (His)<sub>6</sub>-Ub-EGFP encoding human ubiquitin and synthetic EGFP were supplied in a vector (pDG268) by D. A. Gray<sup>143</sup>. PCR (2.2.4.1) was used to amplify DNA inserts to include restriction sites at the 5' and 3' ends of the sequences using Pfu DNA polymerase and the

following primers: Nedd8\_LmrC\_F, Nedd8\_LmrC\_R, Sumo1\_LmrC\_F, Sumo1\_LmrC\_R, EGFPinF and EGFPinR (Table 1). Inserts and vector were then digested using the appropriate restriction enzymes, *NdeI* and *BamHI* to make the NEDD8 and SUMO1 constructs and *NcoI* with *HindIII* to make the EGFP construct (2.2.7). The purified (2.2.5.1) digested DNAs were then ligated (2.2.9) into the corresponding restriction sites of the modified pHUE vector (Fig.2.1).

Screening for successful (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-SUMO1-LmrC and (His)<sub>6</sub>-Ub-EGFP constructs: DH5 $\alpha$ , were transformed (2.3.1), with ligated vector DNA. The resulting ampicillin resistant colonies were then used to inoculate a 10 ml culture (100  $\mu$ g/ml Ampicillin) grown over night and plasmid purified. The plasmid was double digested as described above and run on a 1% (w/v) agarose gel 70 V, to check for bands corresponding to inserts. Successful constructs were then sequenced.

## **2.6.2 Cloning recombinant EF-Tu (GST-TufA and GST-TufB) and GST-elaD expression vectors**

### **2.6.2.1 Cloning TufA and TufB into pGEX4T1**

The *E. coli* elongation factor, EF-Tu is coded for by two genes, *TufA* (GI:147897, protein\_id="AAA50993.1) and *TufB* (GI:297394, protein\_id="CAA40370.1), both of the resulting proteins are identical apart from the final C-terminal amino acid, which is glycine (G) when



coded for by *TufA* and serine (S) when coded for by *TufB*. The cDNA for *TufA* was amplified by PCR using primers FA and RA (Table 1) to include a 5'-*BamHI* site and a 3'-*Sall* site, respectively. The cDNA for *TufB* was amplified by PCR using primers FB and RB which again, include a 5'-*BamHI* site and a 3'-*Sall* site, respectively. The DNA template used to amplify *TufA* and *TufB* was chromosomal DNA obtained from *E. coli* substrain XL10Gold (2.2.3) using PCR mix and cycling parameters (2.2.4.1). *TufA*, *TufB* and pGEX4T1 vector (Stratagene) were then digested using *BamHI* and *Sall* (2.2.7) and ligated (2.2.9) into the *BamHI* and *Sall* sites of pGEX4T1 (Fig.2.2). Ligated vector DNA was then used to transform DH5 $\alpha$  (2.3.1) and resulting ampicillin resistant colonies were screened for the generation of EF-Tu proteins. Successful constructs were then sequenced.

#### 2.6.2.2 Cloning *elaD* into pGEX4T1

The cDNA for *elaD* was amplified by PCR using primers *elaDIF* and *elaDIR* (Table 2.1) to include a 5'-*BamHI* site and a 3'-*NotI* site, respectively from chromosomal DNA obtained from *E. coli* substrain Rosetta™2(DE3) (2.2.3) using PCR mix and cycling parameters (2.2.4.1). *elaD* and pGEX4T1 vector (Stratagene) were then digested using *BamHI* and *NotI* (2.2.7) and ligated (2.2.9) into the *BamHI* and *NotI* sites of pGEX4T1 (Fig.2.2). The ligated vector was then used to transform *E. coli* substrain XL10Gold and resulting ampicillin resistant

colonies were screened for the generation of GST-elaD proteins. Successful constructs were then sequenced.

### **2.6.3 SDM to generate (His)<sub>6</sub>-Ub-LmrC mutants: GA, GP and GGP**

Mutants of the (His)<sub>6</sub>-Ub-LmrC construct: GA, GP and GGP were generated by carrying out PCR based SDM (2.2.4.1) using the methods described in the QuickChange Instruction Manual, Revision A.01 (Stratagene). PCR was carried out using mutagenic primers designed to create mutations in the region linking ubiquitin with LmrC in (His)<sub>6</sub>-Ub-LmrC within the pHUE vector (Fig.2.1). Screening for GA, GP and GGP carrying colonies was carried out first, by transforming DH5 $\alpha$ , (2.3.1) with DpnI digested (2.2.5.2) SDM DNA product. Two of the resulting ampicillin resistant colonies were then used to inoculate a 10 ml culture (100  $\mu$ g/ml Ampicillin), grown over night and purified plasmid was sequenced.

### **2.6.4 SDM to generate GST-elaD mutants: C313S, N227A, W232A, D169A**

Mutants of GST-elaD within the pGEX4T1 vector (Fig.2.2): C313S, N227A, W232A and D169A were generated by carrying out PCR based SDM using Pfu DNA polymerase (Promega) as described in 2.2.4.1

with the exception that a 7 min extension time was used followed by a final 7 min extension and 18 cycles were carried out. Methods described in the QuickChange Instruction Manual, Revision A.01 (Stratagene) were used. PCR was carried out using mutagenic primers designed to create the mutations of *elaD* within the pHUE vector. Screening for mutant carrying colonies was carried out by transforming XL10Gold (2.3.1) with DpnI digested (2.2.5.2) SDM DNA product. Resulting ampicillin resistant colonies were then used to inoculate a 10 ml culture (100 µg/ml Ampicillin), grown over night and purified plasmid was sequenced.

Primer	Primer sequence with restriction sites underlined	Source
Nedd8_LmrC_F	5'-GATGGGCAGCGGCC <u>CATATG</u> CTAATTAAAG-3' ( <i>NdeI</i> )	SIGMA-ALDRICH
Nedd8_LmrC_R	5'-CTGCCTAAGACCGGATCCTCCTCTAAGAG-3' ( <i>Bam</i> HI)	SIGMA-ALDRICH
Sumo1_LmrC_F	5'-GATGGGCAGCGGCC <u>CATATG</u> TCTGACCAG-3' ( <i>NdeI</i> )	SIGMA-ALDRICH
Sumo1_LmrC_R	5'-CTAAACTGTTGAGGATCCCCCGTTTGTTCC-3' ( <i>Bam</i> HI)	SIGMA-ALDRICH
EGFPinF	5'-GGCCGCCCATGGATGAGAGGCAGCCACCA CCATCAC-3' ( <i>NcoI</i> )	SIGMA-ALDRICH
EGFPinR	5'-TGTTGAAAGCTTACTTGTACAGCTCGTCCAT GCCGAG-3' ( <i>HindIII</i> )	SIGMA-ALDRICH
ElaD IF	5'-GCAACTAAAAGGGATCCATGATGGTTACAG-3' ( <i>Bam</i> HI)	SIGMA-ALDRICH
ElaD IR	3'-CGAATTGATTGGGCGGCCGCTTAACTCACT CTTTTGC-5' ( <i>NotI</i> )	SIGMA-ALDRICH
FA	5'-GCACAATAGTAAGGAATGGATCCGTGTCTA AAG-3' ( <i>Bam</i> HI)	'in house'
RA	3'-CTGACTAACGTGGT <u>CGAC</u> GCCCAGAACTTT-5' ( <i>Sall</i> )	'in house'
FB	5'-TCTTAGAGGGAGGATCCATGTCTAAAGAAAA GTTTG-3' ( <i>Bam</i> HI)	'in house'
RB	3'-AAATGTTATCGGGT <u>CGAC</u> GCGCAGAACTTT-5' ( <i>Sall</i> ).	'in house'
ElaDKOF	5'-ATGATGGTTACAGTTGTCAGCAATTATTGTT CATGTGCAGCTCCATCAGCA-3'	SIGMA-ALDRICH
ElaDKOR2	5'-TTAACTCACTCTTTTGCCGGATGCTCTATTC CGCCCAGATACAGAAAAGCC-3'	SIGMA-ALDRICH
ElaDCHF	5'-CAGCACAACCTTTGCAACTAAAAGGAGCCC-3'	SIGMA-ALDRICH
ElaDCHR	5'-GCTGACTGGTCAATTAATTAACAATGATG-3'	SIGMA-ALDRICH

**Table 2.1** Oligonucleotides used in this study with suppliers indicated in the table.

Contents	Final Conc.
dNTPs (10 mM)	200 $\mu$ M of each
5' Primer (41.6 ng/ $\mu$ l)	2.5 ng/ $\mu$ l
3' Primer (41.6 ng/ $\mu$ l)	2.5 ng/ $\mu$ l
Template (15.28 ng/ $\mu$ l)	0.8 – 2.0 ng/ $\mu$ l
Pfu DNA polymerase	0.02 U/ $\mu$ l

**Table 2.2** PCR mix using Pfu DNA polymerase (Promega) were carried out in the PCR components shown in the table, within Pfu buffer (20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 0.1 % (v/v) Triton X-100 and 0.1 mg/ml nuclease-free BSA). DNA was amplified in a thermal cycler. The PCR products were then purified up using QIAquick PCR Purification Kit (QIAGEN) or gel purified using QIAquick Gel Extraction kit (QIAGEN). The resulting cleaned up and purified DNA inserts were then ligated with the corresponding vector.

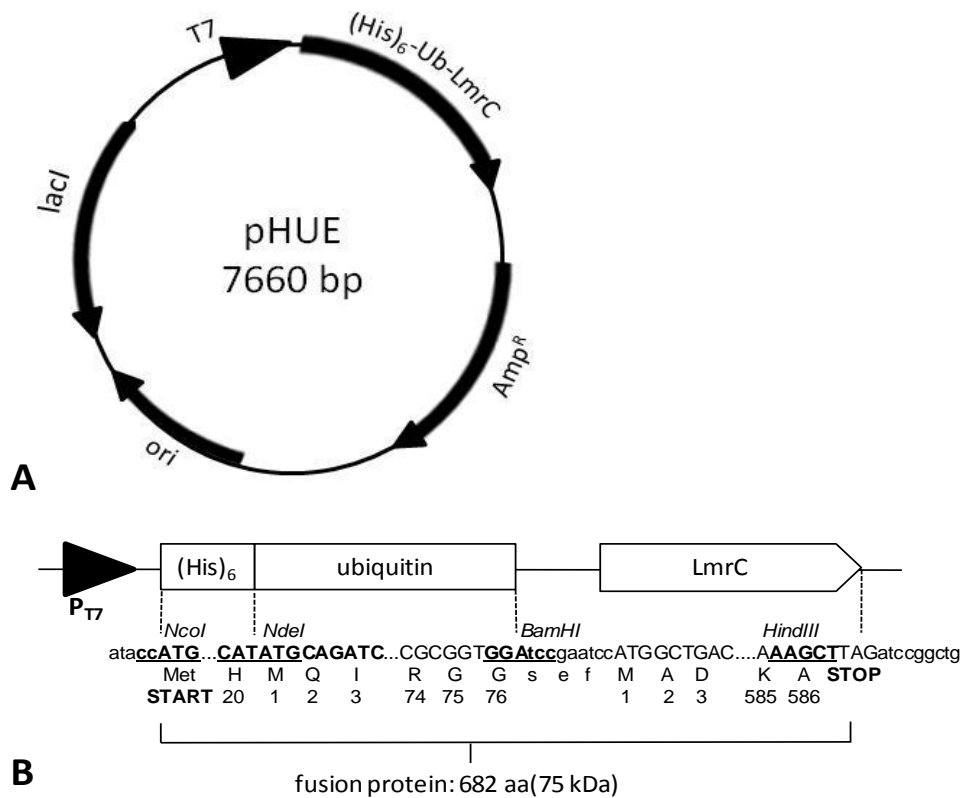
Cycle Step	Temp.	Time	No. of Cycles
Initial denaturation	95 °C	2.5 min	1
Denaturation	94 °C	1 min	30-32
Annealing	*55 °C	1 min	
Extension	72 °C	2 min	
Final extension	72 °C	5 min	1

**Table 2.3** PCR amplification conditions used to amplify DNA using either Pfu DNA polymerase (Promega) or GoTaq<sup>®</sup> DNA Polymerase (Promega). PCR mixes were then purified using QIAquick PCR Purification Kit (QIAGEN) or gel purified using QIAquick Gel Extraction kit (QIAGEN).

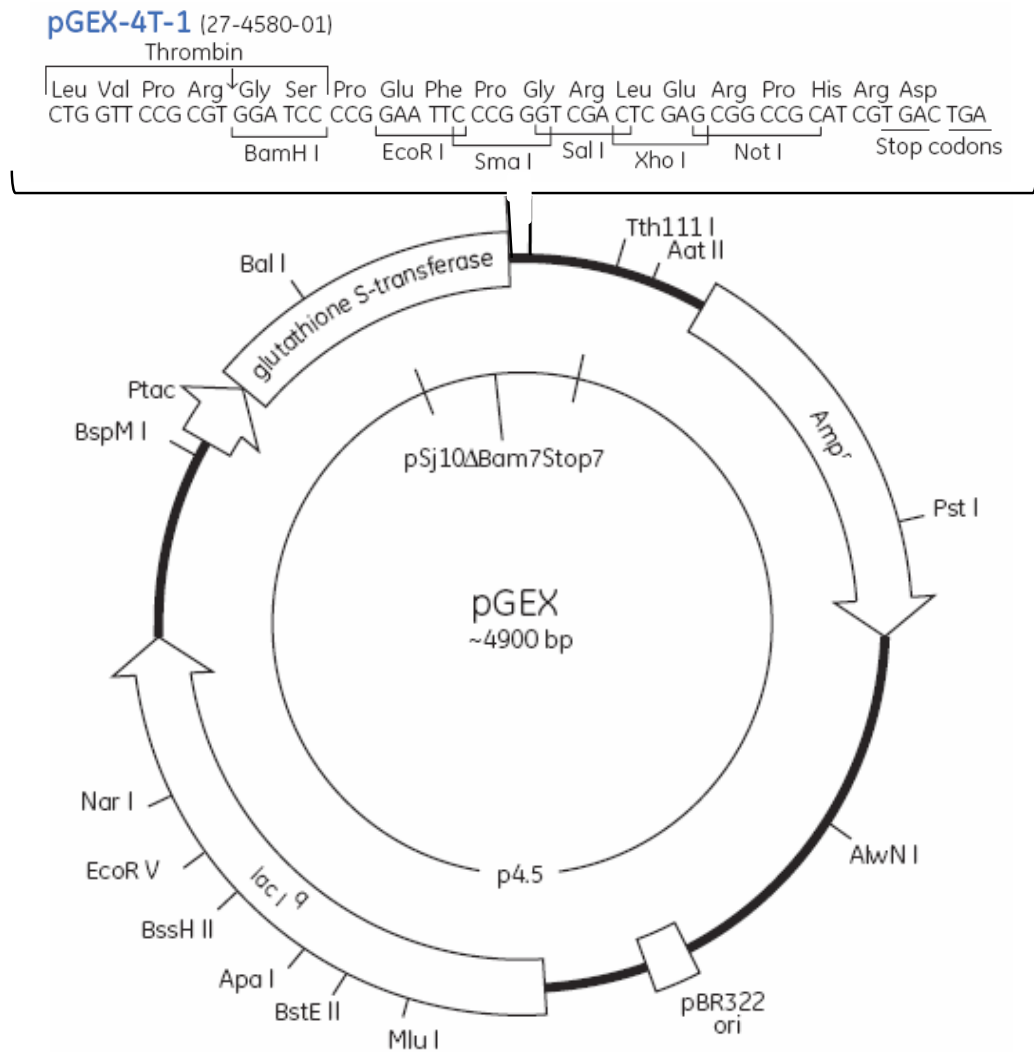
\*When amplifying apramycin resistant template to create the *elaD* knock-out in K12 the annealing temperature = 48°C.

Cycle Step	Temp.	Time	No. of Cycles
Initial denaturation	95 °C	30 s	1
Denaturation	95 °C	30 s	14-22
Annealing	55 °C	1 min	
Extension	68 °C	8 min	

**Table 2.4** The site directed mutagenesis (SDM) cycling parameters used to generate DNA mutants: GA, GP and GGP within the pHUE vector and elaD mutants within the pGEX4T1 vector. This PCR based SDM used Pfu DNA polymerase and methods described in the QuickChange Instruction Manual, Revision A.01



**Fig.2.1.** The pHue vector derived from pET15b (Novagen) was modified by Catanzariti A.M.<sup>4</sup> to encode the T7 promotor in fusion with N-terminal hexahistidine (His)<sub>6</sub> and human ubiquitin (Ub) UBA52, shown in (A). The vector had been further modified by Kerr I.D. (details not provided) to include lincomycin resistance C protein (LmrC) (from the bacterium, *Streptomyces lincolnensis*) which followed immediately after ubiquitin. Specificity of the DUB-like activity in *E. coli* for different fusion proteins was investigated by changing the ubiquitin like protein (UBL) for the UBLs, NEDD8 or SUMO1. Additionally the fusion partner, LmrC was replaced by enhanced green fluorescent protein. The generation of these three alternative fusion proteins was carried out using the region of the pHUE vector, shown in (B). This region of DNA was also used to compare the catalytic mechanism of the DUB-like activity in *E. coli* with those of eukaryotic DUBs, by introducing mutations at -Ub<sup>G76</sup>-S-.



**Fig.2.2.** The empty pGEX4T1 vector, 4969 bp (GE Healthcare) used to generate glutathione S-transferase (GST) tagged proteins with the cloning region available, as indicated. GST-tagged DUB candidates, TufA and TufB were cloned into the pGEX4T1 vector using restriction sites, 5' *BamHI* and 3' *SalI* and *elaD* was cloned in using 5' *BamHI* and 3' *NotI* restriction sites.



### **3 Characterisation of a novel ubiquitin-fusion processing activity in the *E. coli* substrain**

#### **Rosetta™2(DE3)**

##### **3.1 Introduction**

Ubiquitin or ubiquitin-like (UBL) tags have many roles including regulation of protein quality control, function and localisation. Knowledge of the systems for controlling the quality of proteins expressed in prokaryotes is limited and in bacteria, this does not involve ubiquitin. It is known that archaea express ubiquitin<sup>26</sup>, but bacteria (eubacteria) do not. However, there have been a limited number of reports observing intrinsic, low level DUB-like activity in *E. coli* expressing ubiquitin fusion proteins. It is common practice to express a protein as part of a fusion protein with a UBL (SUMO1 or ubiquitin) as the UBL improves folding of the nascent peptide and maintains solubility when expressed in bacteria. It is from studies using these fusion proteins that observations have been made which conflict with the assumption at the time of starting this project (2005), that as bacteria do not express ubiquitin, they would not express enzymes which specifically cleave a ubiquitin fusion protein.

One group has reported DUB-like activity in *E. coli* on two occasions. In 1987 *E. coli* substrain AR58 was used to express linear, penta-ubiquitin equivalent to a ubiquitin precursor and immunoblotted for ubiquitin, which revealed the break down products from penta-ubiquitin: mono-, di-, tri- and tetra-ubiquitin<sup>3</sup>. Two years later, while developing a robust ubiquitin fusion protein expression system, again in *E. coli* substrain AR58, the group observed low level DUB-like activity<sup>138</sup>.

On another occasion DUB-like activity was not only observed in *E. coli* substrain BL21(DE3) but also demonstrated to result in specific cleavage of a substrate<sup>4</sup>. DUBs cleave immediately after glycine (G) of ubiquitin (Ub) of the C-terminus (Ub<sup>G76</sup>). In this case the authors expressed ubiquitin, tagged with polyhistidine, (His)<sub>6</sub> residues at the N-terminus, fused to a broad range of proteins in differing sizes to develop a system for high level recombinant protein expression. These included two mutants of glutathione S-transferase with M or P at the N-terminus (M-GSTP1, P-GSTP1 respectively). Hence, the fusion proteins expressed were; (His)<sub>6</sub>Ub-SUMO; (His)<sub>6</sub>Ub-M-GSTP1 and (His)<sub>6</sub>Ub-P-GSTP1. Interestingly, all the fusion proteins investigated, when expressed in *E. coli* substrain BL21(DE3) were cleaved, except for Ub-P-GSTP1. The authors confirmed this was a specific activity by protein sequencing the cleavage products from (His)<sub>6</sub>Ub-SUMO and (His)<sub>6</sub>Ub-M-GSTP1. The specific DUB-like activity, as well as the resistance of

Ub-P to cleave implicates UCH activity. However, two homologues of ubiquitin specific enzymes, Unp (mouse) and Unph (human) have been reported to cleave a Ub-P bond, verified by N-terminal sequencing, suggesting the existence of an alternative DUB mechanism <sup>144</sup>.

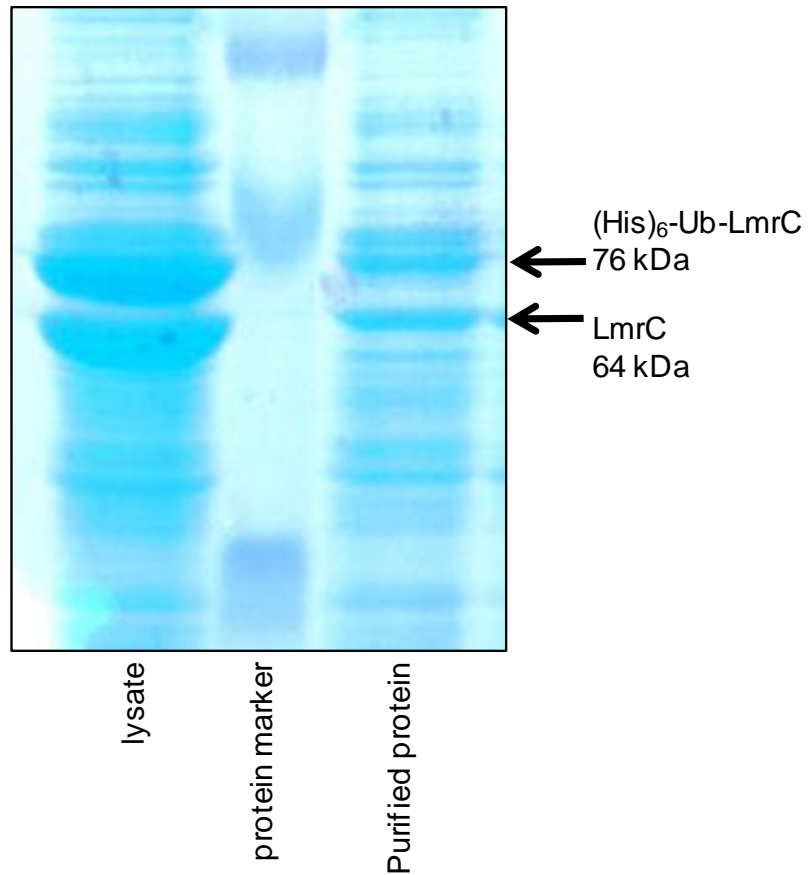
Further insights to DUB activity in *E. coli* were gleaned from *in vitro* work carried out by A. Ciechanover and colleagues<sup>139</sup>, while characterising a protein termed factor Hedva (FH). In this case the substrate was a core nucleosomal histone (H2A) conjugated to ubiquitin via an isopeptide bond, rather than a peptide bond, as in the cases mentioned above. The authors reported the inhibition of the DUB-like activity from *E. coli* substrain DH5 $\alpha$  in the presence of a UCH inhibitor, ubiquitin aldehyde, suggesting that the enzyme is a UCH (one which can also cleave isopeptide bonds). Of particular note, this group also reported the co-migration of an unexpected candidate for this activity, the bacterial elongation factor, EF-Tu. The human elongation factor, EF-1 $\alpha$  was also proposed to harbour the same activity. The structure and function of elongation factors in the delivery of charged tRNA to ribosomes during translation has been fully described <sup>140-142</sup>.

### **3.2 Work leading up to this study**

The starting point for this work was initiated by the observation of a ubiquitin-fusion processing activity attributed to an undocumented expression strain of *E. coli* known to be a DE3 strain, by other members of this group (H. Jefferey and I. Kerr). Those members were investigating the mechanism of antibiotic-resistance-conferring ATP-binding cassette (ABC) proteins. ABC proteins use the hydrolysis of bound ATP to undergo a conformational change which drives the transport of ions and macromolecules across the cell membrane. Some ABC proteins are associated with antibiotic resistance by pumping antibiotics out of the cell. ABC proteins are typically made up of 4 domains: two transmembrane domains (TMD) which span the membrane bilayer and two ABCs also known as nucleotide binding domains (NBDs), placed in the cytoplasm<sup>145-147</sup>. The ABC protein under investigation was lincomycin resistance C (LmrC), a twin-NBD protein (60 kDa) which lacked TMDs. LmrC is involved in resistance to the antibiotic, lincomycin; both LmrC and lincomycin are expressed by the bacterium, *Streptomyces lincolnensis*.

LmrC was over-expressed in a DE3 strain of *E. coli* as the fusion protein, (His)<sub>6</sub>-Ub-LmrC, with an N-terminal hexahistidine (His)<sub>6</sub>, by H. Jefferey and I. Kerr. The intention was to improve the recombinant protein expression level. The expressed (His)<sub>6</sub>-Ub-LmrC protein was

purified by Ni<sup>2+</sup>-affinity chromatography then resolved using 10 % (v/v) SDS-PAGE and stained with Coomassie blue (Fig.3.1). Just below the 75 kDa band representing the intact fusion protein, a 64 kDa band was noted, which was presumed to represent LmrC. Although this cleaved LmrC would not have a (His)<sub>6</sub> tag for capture during protein purification, it is possible that the intact (unprocessed) fusion protein can form a dimer with the LmrC cleavage product and so could be co-purified. The presence of a 64 kDa band posed the question as to whether it was specific DUB-like cleavage. This discovery and the observations of others in the literature indicated that there may be an inherent DUB-like activity in *E. coli* or in some of its substrains. Collectively it was these findings which initiated the current study, to characterise the ubiquitin-fusion processing activity in *E. coli* initially using substrain Rosetta™2(DE3).



**Fig.3.1** A substrain of *E. coli*, known to be a DE3 strain carrying a pHUE vector, sub-cloned with (His)<sub>6</sub>-Ub-LmrC was grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (50 µM) was used to induce protein expression. The resulting protein was purified using nickel affinity chromatography, electrophoresed on a 10% (w/v) acrylamide gel and stained using Coomassie Blue stain. A 76 kDa band for (His)<sub>6</sub>-Ub-LmrC and a lower band representing the 64 kDa cleavage product, LmrC was noted, confirmed using N-terminal sequence analysis. This led to the proposal that the cleavage products resulted from DUB cleavage in the DE3 strain of *E. coli* (H. Jefferey and I. Kerr, unpublished). These observations, together with those in the literature led to the current study, to characterise the ubiquitin-fusion processing activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3).

### **3.3 The aims and objectives of the study**

The aims of this study were fourfold. The initial aim was to demonstrate the reproducibility of the observed cleavage of the (His)<sub>6</sub>-Ub-LmrC fusion protein when expressed in *E. coli* substrain Rosetta™2(DE3). Then, the nature of the type of cleavage, specific DUB-like or non-specific had to be determined. Thirdly, once the type of cleavage was known, the specificity of the ubiquitin-fusion processing activity expressed in *E. coli* substrain Rosetta™2(DE3) was then investigated. Finally, a candidate approach was taken to identify the enzyme or enzymes responsible for cleaving the (His)<sub>6</sub>-Ub-LmrC fusion protein when expressed in *E. coli* substrain Rosetta™2(DE3). In this chapter, the first two aims of the study are explored. The main thrust was to investigate specificity using Western blotting techniques, N-terminal sequencing analysis, mass spectrometry and site directed mutagenesis (SDM).

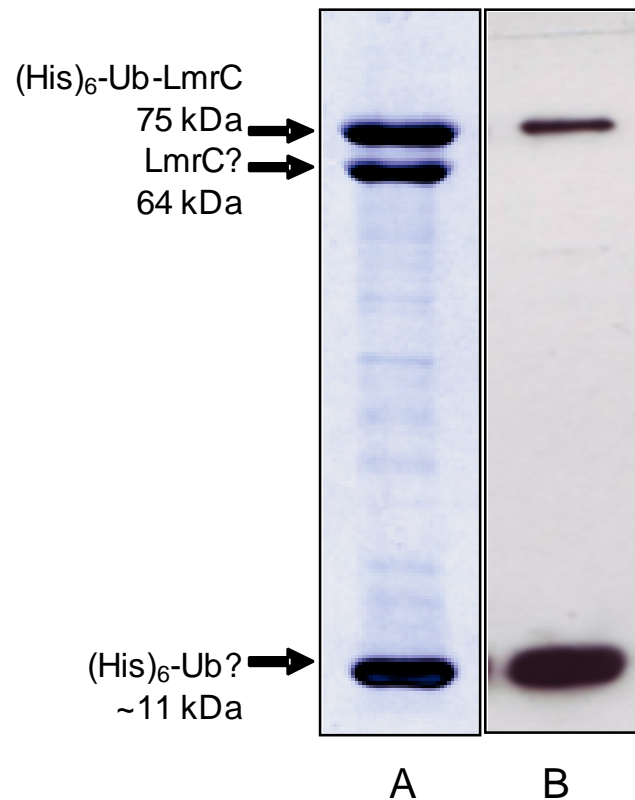
### **3.4 Results**

#### **3.4.1 *In vivo* cleavage of the (His)<sub>6</sub>-Ub-LmrC fusion protein in *E. coli* substrain Rosetta™2(DE3)**

In order to check the reproducibility of the original findings of other members of this group (H. Jefferey and I. Kerr), using the same procedures as carried out before, (His)<sub>6</sub>-Ub-LmrC was expressed in *E. coli* substrain Rosetta™2(DE3), purified using cobalt affinity chromatography and analysed by SDS-PAGE on a gradient gel, 5-20% (w/v) acrylamide. Previously, an undocumented expression strain of *E. coli* known to be a DE3 strain, had been used for the expression of recombinant protein, therefore the DE3 strain of *E. coli*, Rosetta™2(DE3) was chosen for this purpose in the current study. The Coomassie stained gel (Fig.3.2 A) has a similar banding pattern when compared with the Coomassie stained gel in Fig.3.1 (the findings of H. Jefferey and I. Kerr). In Fig.3.2 there is a 75 kDa band for intact (His)<sub>6</sub>-Ub-LmrC and a 64 kDa band, presumed to represent LmrC. Therefore it is clear from these results that the original findings of ubiquitin-fusion processing activity in a DE3 strain in *E. coli*, were reproducible. In support of these results there was a 75 kDa, immunoreactive band representing the intact fusion protein when using the ubiquitin antibody, yet no immunoreactive band for LmrC (64 kDa) (Fig.3.2.B), which was to be expected given that (His)<sub>6</sub>-Ub would have been removed from



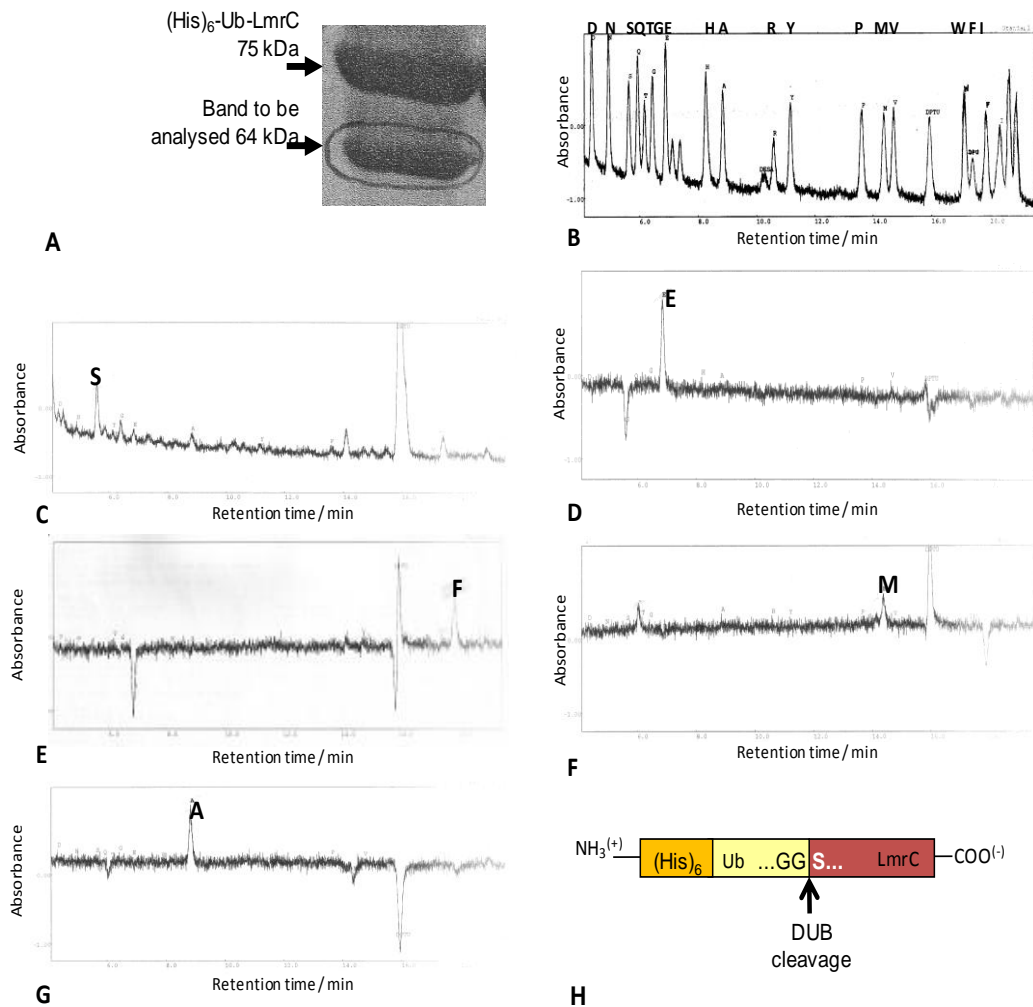
LmrC during cleavage. Additionally, these findings also show an 11 kDa band which was immunoreactive when probing for ubiquitin, so could possibly represent 'free' (His)<sub>6</sub>-Ub.



**Fig.3.2** *E. coli*, Rosetta<sup>TM</sup>2(DE3) cells carrying a pHUE vector, sub-cloned with  $(\text{His})_6\text{-Ub-LmrC}$  were grown in LB plus ampicillin (100  $\mu\text{g/ml}$ ) until optical density = 0.6-0.8. IPTG (50  $\mu\text{M}$ ) was used to induce protein expression. The resulting protein was purified using cobalt affinity chromatography, electrophoresed on a gradient gel 5-20% (w/v) acrylamide and either stained using Coomassie Blue stain (A) or transferred on to nitrocellulose membrane and probed for ubiquitin (B). It is clear from these results that the original findings of a ubiquitin-fusion processing activity in a DE3 strain in *E. coli* were reproducible in Rosetta<sup>TM</sup>2(DE3). A 75 kDa band was evident in both the Coomassie stained gel (A) and the Western blot (B). The 64 kDa band presumed to represent LmrC was present in the Coomassie stained gel, but not when probed for ubiquitin, which was to be expected given that  $(\text{His})_6\text{-Ub}$  would have been removed from LmrC during cleavage. In addition to providing evidence that this cleavage in *E. coli* was reproducible, these results also show a smaller immunoreactive band (~11 kDa) which could represent the cleavage product,  $(\text{His})_6\text{-Ub}$  (B).

### **3.4.2 Investigating the nature of the cleavage of the (His)<sub>6</sub>-Ub-LmrC fusion protein**

To test the proposals in section 3.4.1, that the ubiquitin-fusion processing activity was a DUB activity, it was necessary to sequence the N-terminus of the 64 kDa protein observed in Fig.3.2.A to determine precisely where in the fusion protein cleavage had occurred. This would inform about the specificity of the enzyme responsible. This was done using a form of N-terminal sequence analysis, termed Edman degradation. Purified proteins were separated on a 5-20% (w/v) gradient gel and transferred on to Polyvinylidene fluoride (PVDF) membrane and stained with Amido Black solution (Fig.3.3). The preparation of purified proteins for N-terminal sequence analysis is detailed in 2.4.6.1. The circled band (~64 kDa), was submitted for N-terminal sequence analysis. The N-terminal sequence analysis results show a specific DUB cleavage of the fusion protein had occurred, immediately after G76 of ubiquitin; the N-terminal sequence of the first five amino acids of the 64 kDa protein was found to be SEFMA, which is the linker sequence for LmrC that immediately follows G76 of ubiquitin. This indicated that a specific DUB cleavage had occurred in *E. coli* substrain Rosetta™2(DE3).

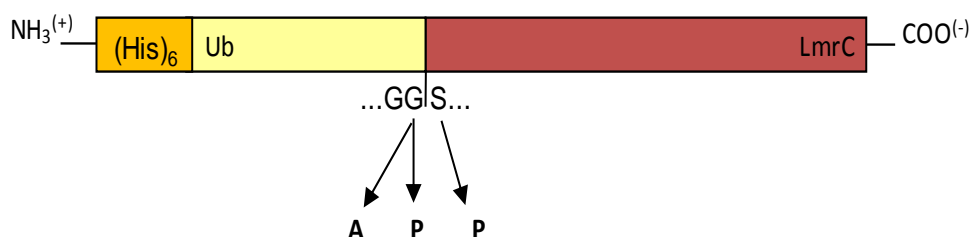


**Fig.3.3.** Purified (His)<sub>6</sub>-Ub-LmrC fusion protein was separated on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and stained with Amido Black solution. The 64 kDa band submitted for N-terminal sequence analysis (Department of Biochemistry, University of Cambridge) is indicated (A). Chromatogram (B) shows retention times (min) of derivatised amino acid standards. (C-G) indicate the derivatised amino acids released from the 64 kDa protein to be SEFMA which is the linker sequence to LmrC that immediately follows G76 of ubiquitin within the fusion protein (H). This revealed that a specific DUB cleavage had occurred in *E. coli* Substrain Rosetta<sup>TM</sup>2(DE3) after G76 of ubiquitin.

### **3.4.3 Characterising the specificity of the DUB activity in *E. coli* substrain Rosetta™2(DE3) by modifying residues around the cleavage site in the (His)<sub>6</sub>-Ub-LmrC substrate**

In order to determine if the DUB activity in *E. coli* displayed similar characteristics to those reported for eukaryotic DUBs, it was necessary to generate mutants of the cleavage site of ubiquitin, Ub<sup>G76</sup>x within the (His)<sub>6</sub>-Ub-LmrC fusion protein. Proline (P) as a general rule, is known to render fusion proteins refractory to cleavage if in position 'x' of a DUB site, when expressed in bacteria, yeast or mammals<sup>5</sup>. There is one exception, the Unp enzyme and homologue, Unph<sup>144</sup> which is thought to operate under a different catalytic mechanism. Therefore, activity towards the Ub-P bond could shed light on the catalytic nature of the prokaryotic DUB activity. Proline has also been shown to prevent cleavage of fusion proteins in studies using a fusion protein containing a ubiquitin mutant, Ub<sup>G76P</sup>. One such study, expressed (His)<sub>6</sub>-Ub fused with enhanced green fluorescent protein (EGFP), (His)<sub>6</sub>-Ub<sup>G76P</sup>-EGFP in mammalian cells (D.A. Gray personal communication). A fusion protein containing a Ub<sup>G76A</sup> mutant, (His)<sub>6</sub>-Ub<sup>G76A</sup>-EGFP was also used in the same study and was shown to offer no resistance to cleavage in mammalian cells. In contrast however, a Ub<sup>G76A</sup> mutant, when fused to metallothionein (MT) in the fusion protein, Ub<sup>G76A</sup>-MT perturbed cleavage when expressed in yeast<sup>148</sup>.

In the current study the DNA coding for (His)<sub>6</sub>-Ub-LmrC was mutated in the region between ubiquitin and LmrC to determine how the catalytic mechanism of the ubiquitin-fusion processing activity in *E. coli* substrain Rosetta™2(DE3) compared with mechanisms of eukaryotic DUBs. DUB cleavage of (His)<sub>6</sub>-Ub-LmrC would result in two cleavage products, (His)<sub>6</sub>-Ub and LmrC. The final residues (75 and 76) of ubiquitin are GG and the DUB site in (His)<sub>6</sub>-Ub-LmrC follows the last amino acid (76<sup>th</sup>) C-terminal to ubiquitin (Ub<sup>G76</sup>) prior to a serine (S) in a linker at the start of the LmrC protein (Fig.3.4). Therefore (His)<sub>6</sub>-Ub-LmrC can be denoted as follows, Ub<sup>G76</sup>-S-LmrC. Two mutants were made by converting Ub<sup>G76</sup> within (His)<sub>6</sub>-Ub<sup>G76</sup>-S-LmrC to either alanine (A) or proline (P) to produce mutants, GA or GP respectively. A third mutant (mutant GGP) was made by converting the S in (His)<sub>6</sub>-Ub<sup>G76</sup>-S-LmrC, immediately following Ub<sup>G76</sup>, to P. These three (His)<sub>6</sub>-Ub-LmrC mutants will subsequently be referred to a GA, GP and GGP throughout the study (Fig.3.4). GA, GP and GGP were generated by carrying out PCR based SDM using Pfu DNA polymerase as described in 2.6.3.

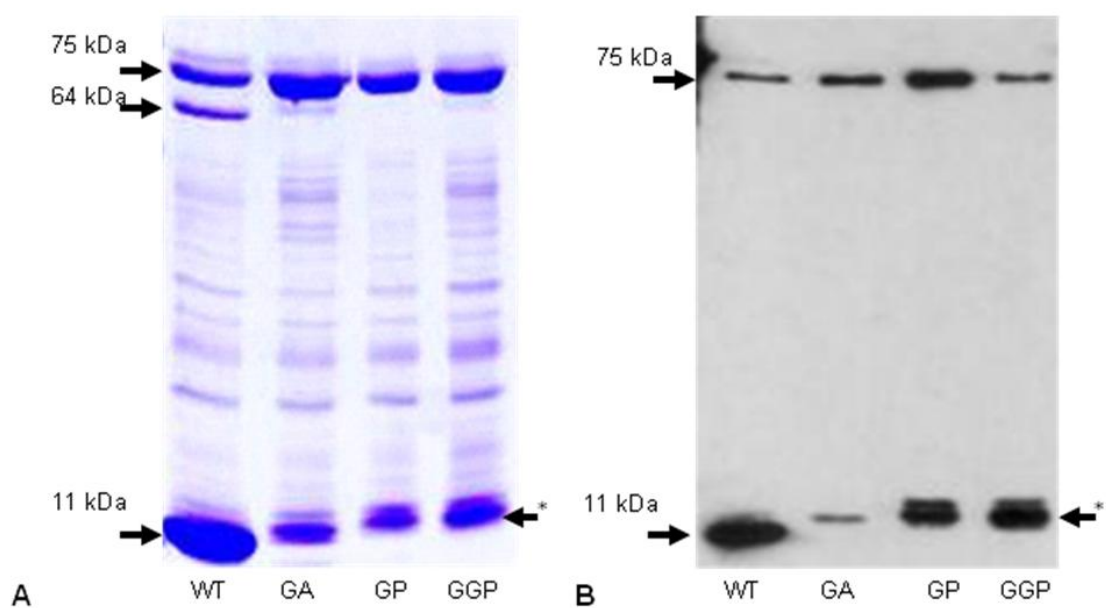


**Fig.3.4 The fusion protein, (His)<sub>6</sub>-Ub-LmrC** consists of hexahistidine (His)<sub>6</sub>, human ubiquitin (Ub) UBA52 and the full sequence for lincomycin resistance C protein (LmrC) from the bacterium, *Streptomyces lincolnensis*. The DNA coding for (His)<sub>6</sub>-Ub-LmrC was mutated at the DUB cleavage site to compare catalytic mechanisms of the DUB-like activity in *E. coli* with mechanisms of eukaryotic DUBs. The cartoon shows three amino acids within (His)<sub>6</sub>-Ub-LmrC which were each mutated to produce three different mutants. The 76<sup>th</sup> residue of ubiquitin, glycine (G) was mutated to either alanine (A) to produce the first mutant, GA, or to proline (P) to create the second mutant, GP. The third mutant, GA was generated by mutating the serine (S) which immediately follows the 76<sup>th</sup> residue of ubiquitin to P.

Wild type (WT) (His)<sub>6</sub>-Ub-LmrC fusion protein and mutants, GA, GP and GGP were expressed in *E. coli* substrain Rosetta™2(DE3), purified and analysed by SDS-PAGE on a gradient gel, 5-20% (w/v) acrylamide (Fig. 3.5) and either stained using Coomassie Blue stain (Fig.3.5.A) or transferred on to nitrocellulose membrane and immuno-probed for ubiquitin (Fig.3.5.B). A 75 kDa band representing the intact fusion proteins was observed for WT and mutants on the Coomassie stained gel and the blot. The 64 kDa cleavage product observed in the WT sample was not seen in the mutants, however a similar band was evident for the mutants but less abundant, particularly for GA, with a larger, denatured molecular weight, evident for all three mutants. A band corresponding to the smaller DUB cleavage product was only present for the WT fusion protein, and a band with a slightly higher mass than this was present for all mutants. Without knowing the identity of these proteins it was not possible at this stage to know if they were DUB cleavage products showing aberrant migration, or if the proteins did have a higher mass and were therefore likely to be the products from an alternative, non-specific cleavage site. In order to determine the true nature of the cleavage of the mutant fusion proteins, N-terminal sequencing analysis was used to determine the specific N-terminal sequence of the larger bands produced by the mutants in the 65 kDa region. Additionally, it was necessary to use protein mass spectrometry



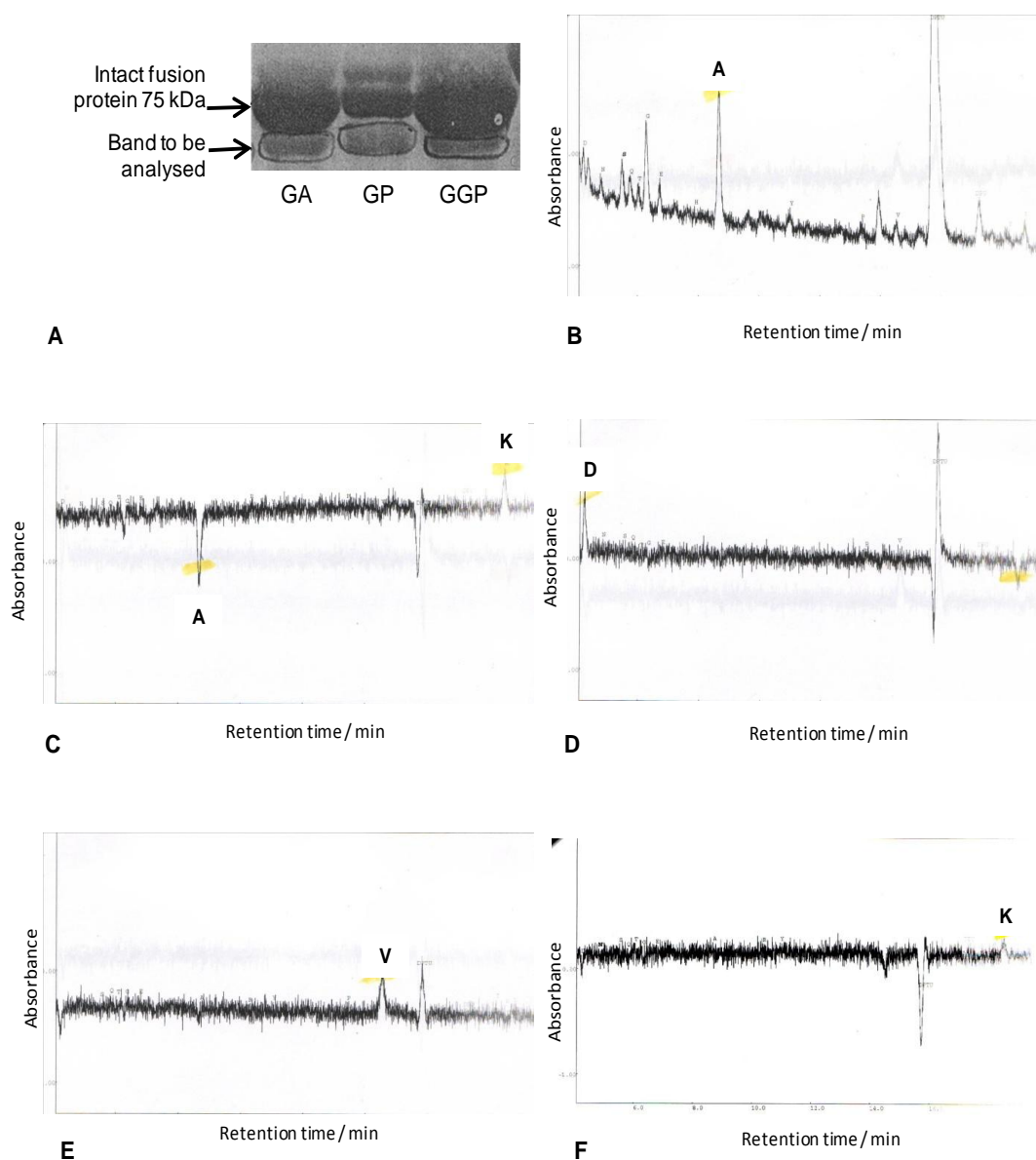
(MS) to determine the exact masses of the smaller cleavage products and by inference, sequences.



**Fig.3.5** *E. coli*, Rosetta™2(DE3) cells carrying a pHUE vector, sub-cloned with WT (His)<sub>6</sub>-Ub-LmrC (WT) or one of three mutants GA, GP and GGP were grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (50 µM) was used to induce protein expression. The resulting protein was purified using cobalt affinity chromatography, electrophoresed on a gradient gel, 5-20% (w/v) acrylamide and either stained using Coomassie Blue stain (A) or transferred on to nitrocellulose membrane and probed for ubiquitin (B). There was a band similar to the 64 kDa protein in the mutants compared with WT (with a lower intensity in the mutants). A DUB-like cleavage product (11 kDa) was present for WT and a protein band (\*) of slightly larger mass than to be expected for DUB cleavage was also seen for all mutants.

#### **3.4.4 Edman degradation to 'pinpoint' the cleavage site of the GA,**

GP and GGP mutant fusion proteins: Purified GA, GP and GGP were separated on a 5-20% (w/v) gradient gel and transferred on to Polyvinylidene fluoride (PVDF) membrane and stained with Amido Black solution (Fig.3.6). The preparation of purified proteins for N-terminal sequence analysis is detailed in 2.4.6.1. The circled bands, initially assumed to be processed LmrC (~64 kDa), previously detected in WT (His)<sub>6</sub>-Ub-LmrC was submitted for N-terminal sequence analysis. The chromatograms in Fig.3.6.B - Fig.3.6.F) indicate the first six amino acids of the band analysed from the GA lane, were AAKDVK which sequence alignments show is identical to that of the 60 kDa chaperone protein, GroEL (Cpn60). GroEL sequence was also present in bands from the GP and GGP lanes, data not shown. Therefore these results do not provided evidence of LmrC sequence in the ~64 kDa band for any of the mutants. It would clearly be informative to determine the molecular weights of the smaller bands of the WT, GA, GP and GGP cleavage product within the ~11 kDa region.

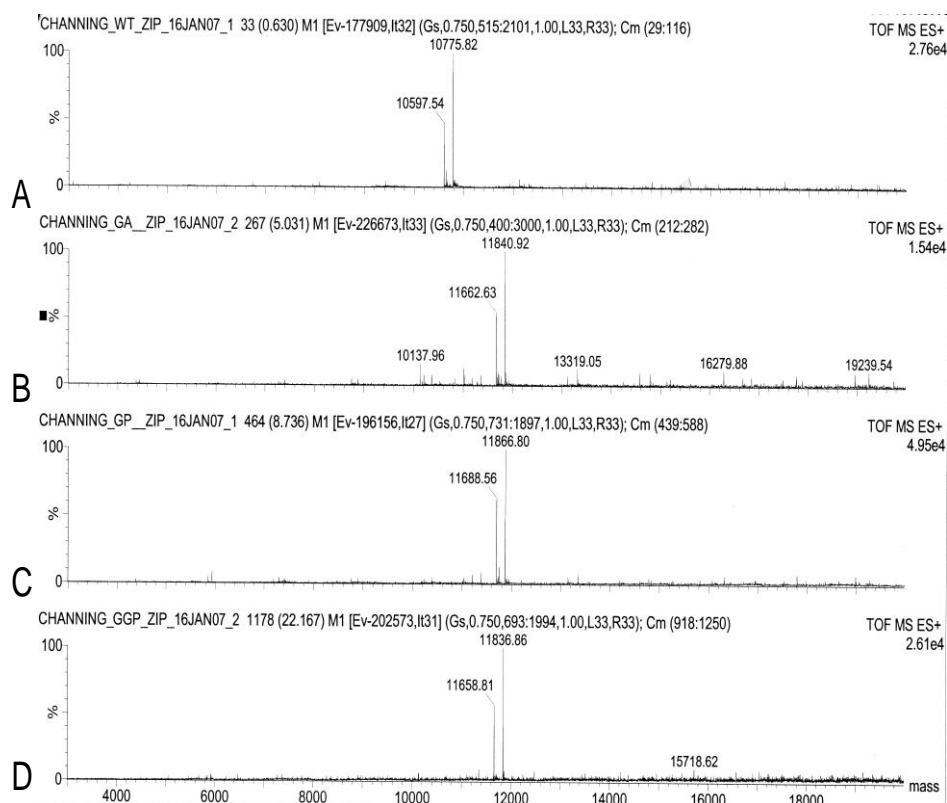


**Fig.3.6** (A) Purified (His)<sub>6</sub>-Ub-LmrC mutant fusion proteins, GA, GP and GGP were separated on a gradient gel 5-20% (w/v) acrylamid and the gel was incubated in glycine free transfer buffer (2.4.6.1), rocking (15 min), transferred on to nitrocellulose membrane and stained with Amido Black. A higher protein load was used this time compared with the proteins loaded in Fig.3.5. The protein bands indicated in (A) were submitted for N-terminal sequence analysis (Department of Biochemistry, University of Cambridge). The resulting chromatograms (B-F) show retention times (min) of derivatised amino acids released from the GA fusion product, which sequence alignments show is identical to that of the 60 kDa chaperone protein, GroEL (Cpn60). GroEL was also present in bands from the GP and GGP samples, data not shown.

### **3.4.5 Mass spectrometry to determine the precise mass of (His)<sub>6</sub>-Ub and the proteins present in the ~11 kDa region**

MS MALDI-TOF analysis was used to find the exact masses of proteins in the ~11 kDa region in purified protein samples expressed by *E. coli* substrain Rosetta<sup>TM</sup>2(DE3), carrying WT, GA, GP or GGP constructs. The resulting protein samples analysed will be referred to as WT, GA, GP and GGP samples, to indicate the construct that was used to generate the protein sample analysed. Fig.3.7A - Fig.3.7D) shows the spectra resulting from WT, GA, GP and GGP protein samples analysed, respectively, mass (Da) is shown on the x-axis and the percentage of maximum peak ion (%) is shown on the y-axis. There are two peaks shown for each sample analysed, data also included in Table 3.1. The estimated masses (Da) for the predicted (His)<sub>6</sub>-tagged DUB cleavage products for all four samples analysed (WT, GA, GP and GGP) are also included in Table 3.1. The masses for theoretical DUB cleavage products, after: Ub<sup>G76</sup>, Ub<sup>A76</sup> and Ub<sup>P76</sup> in WT, GA or GP respectively, in the 11 kDa region (Table 3.2) were estimated for WT and all mutant fusion proteins using the online tool, 'Compute pI/Mw' (ExPASy). Masses were estimated using values corresponding to nitrogen as found in nature ('natural isotopes'). These estimated values include proteins without an N-terminal methionine as proteins can also undergo N-terminal methionine (M) excision (NME), according to the N-End rule<sup>149</sup>. It is evident in Table 3.1 that only the WT sample, analysed had

an estimated value, 10596.98 Da (-M) close to an MS peak value, 10597.6 Da (peak 1) for a DUB cleavage product. This evidence supports the findings from N-terminal sequence analysis, showing the WT fusion protein was cleaved specifically at the DUB site. However, 'peak 1' values for all the mutants are larger (by 1062 Da for GGP and 1052 Da for GA and GP) indicating that these are likely to be products from a cleavage site to the C-terminal of the DUB site within the LmrC sequence.



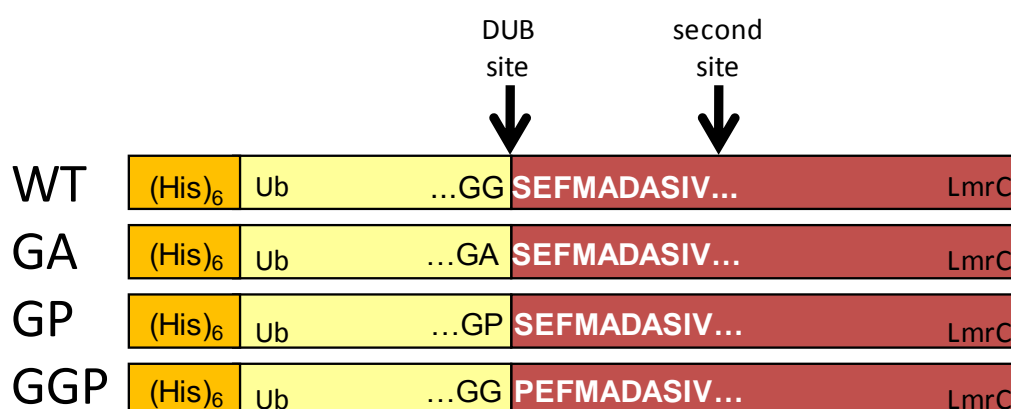
**Fig.3.7** *E. coli*, Rosetta™2(DE3), transformed with (His)<sub>6</sub>-Ub-LmrC (WT) or one of three mutants GA, GP and GGP were grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (50 µM) was used to induce protein expression. The protein was purified using cobalt affinity chromatography then submitted for analysis by MALDI-TOF mass spectrometry (S. Liddell, University of Nottingham) in the ~11 kDa region. The resulting spectra for proteins purified from WT, GA, GP and GGP bacterial cultures are shown in (A – D) respectively, with mass (Da) is shown on the x-axis and the percentage of maximum peak ion (%) is shown on the y-axis. Two peaks and therefore two proteins are shown for each of the four protein samples analysed. Table 3.1 also summarises these results.

Construct	MM of proteins predicted to be generated by precise cleavage after G76, A76 or P76 of indicated proteins (Da)		MS analysis of proteins in the ~11 kDa region (Da)	
	Unmodified (+ M)	Modified (– M)	Peak 1	Peak 2
WT	10728.17	10596.98	10597.6	10775.8
GA	10742.19	10611.00	11662.7	11840.9
GP	10768.23	10637.04	11688.6	11866.8
GGP	10728.17	10596.98	11658.8	11836.8

**Table 3.1** *E. coli*, Rosetta™2(DE3), transformed with (His)<sub>6</sub>-Ub-LmrC (WT) or one of three mutants GA, GP and GGP were grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (50 µM) was used to induce protein expression. The protein was purified using cobalt affinity chromatography then submitted for analysis by MALDI-TOF mass spectrometry, in the ~11 kDa region. There were two peaks for each protein sample submitted for analysis, summarised in the table. The online tool, 'Compute pI/Mw' (ExPASy) was used to estimate molecular masses (MM) of proteins predicted to be generated by precise cleavage after G76, A76 or P76 of indicated proteins (Da). A post-translational modification of proteins, involving the N-terminal excision of methionine (M) can occur as part of the N-end rule, so values for this 'modified' cleavage product (– M) have also been estimated. Masses were estimated using values corresponding with nitrogen as found in nature, 'natural' isotopes. These estimated values for the cleavage product were then compared with those from the MS results (peak 1). It is evident from these data that the estimated mass, 10596.98 Da (-M) from the WT sample, is very close to the mass, 10597.6 Da, from the MS results (peak 1). This suggests accurate DUB cleavage occurred with the WT fusion protein, supporting the findings from N-terminal sequence analysis (Fig.3.3). However, none of the estimated masses are similar to masses from the MS results (peak 1) for either, GA, GP or GGP samples analysed, which would indicate that in this case the cleavage products did not result from specific DUB cleavage.



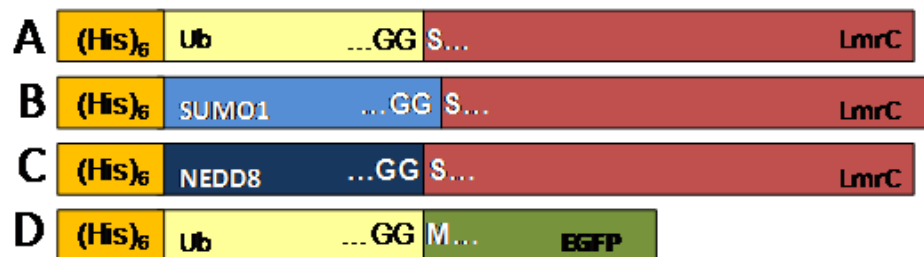
In order to ascertain that the ~11 kDa bands seen in mutant proteins in Fig.3.5, resulted from an alternative cleavage site, the masses of proteins predicted from proteolysis at different positions within the LmrC sequence were calculated. This indicated an alternative cleavage site was located 10 aa C-terminal to the DUB site, in all mutants, (see cartoon in Fig.3.8). The MS data confirmed the WT fusion protein cleaved specifically at the DUB site when expressed in *E. Coli* substrain Rosetta™2(DE3), which supports the evidence obtained by N-terminal sequencing. Therefore only the WT fusion protein was specifically cleaved by the prokaryotic DUB activity, suggesting that this activity uses a 'conventional' catalytic mechanism. The second, larger peak (by 178 Da) observed in Fig. 3.7 for WT and mutant samples analysed by MS was likely to be a modified form of the cleavage product. When expressed in *E. coli* (His)<sub>6</sub>-tagged proteins have been shown to undergo  $\alpha$ -N-6-phosphogluconoylation of hexahistidine which results in the extra mass of 178 Da<sup>150</sup>. At the time of writing, the Merops database was used to search for a peptidase that would recognised the second cleavage site and no peptisae was found.



**Fig.3.8** A cartoon to show two cleavage sites within wild type (His)<sub>6</sub>-Ub-LmrC (WT), GA, GP and GGP fusion proteins, when expressed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). The first cleavage site is a DUB site immediately following Ub<sup>G76</sup> and the second cleavage site is 10 amino acids downstream of the DUB site, within LmrC. WT was shown by N-terminal sequence analysis as well as MS analysis to cleave specifically at the DUB site. In contrast, GA, GP and GGP were all demonstrated, using MS analysis to cleave non specifically, at the second site. The absence of DUB cleavage of GA, GP and GGP by the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) suggests that the DUB activity uses a 'conventional' catalytic mechanism. At the time of writing, the Merops database did not have a peptidase which recognised the second cleavage site.

#### **3.4.6 Further characterising of the DUB activity in *E. coli* substrain Rosetta™2(DE3) by replacing either one of the fusion proteins within the (His)<sub>6</sub>-Ub-LmrC fusion protein**

Having revealed a conventional catalytic mechanism in the DUB activity from *E. coli* substrain Rosetta™2(DE3) the next step was to change the fusion partners within the fusion protein to investigate any preferences of the DUB activity for a particular fusion protein. This could potentially shed light on the type of substrate that this prokaryotic DUB activity normally targets. The UBLs, SUMO1 or NEDD8 were used to replace ubiquitin, and EGFP was used to replace LmrC (Fig. 3.9).



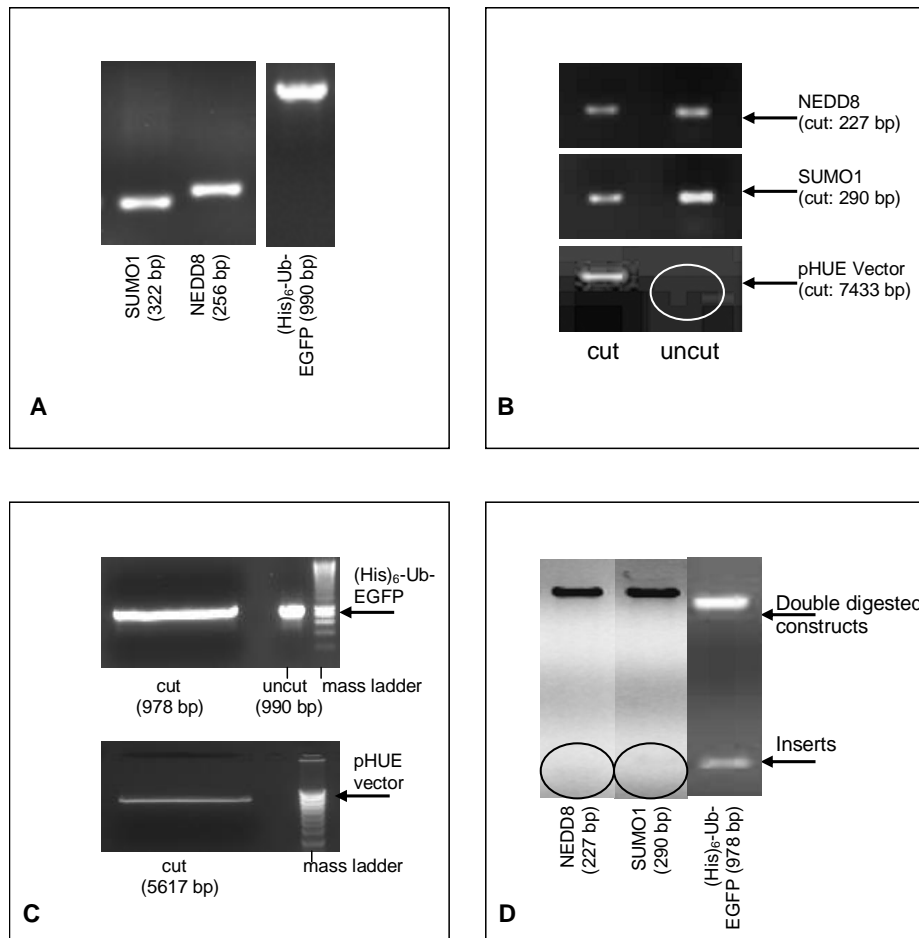
**Fig.3.9** Cartoon to show the changes made to (His)<sub>6</sub>-Ub-LmrC (A) to construct three alternative fusion proteins (B-D). The changes were made to the region of DNA encoding (His)<sub>6</sub>-Ub-LmrC in the pHUE vector. Ub was replaced with either SUMO1 (B) or with NEDD8 (C). To construct the third fusion protein (D), LmrC was replaced with enhanced green fluorescent protein (EGFP). This was performed to investigate any differences in specificity that the DUB activity in *E. coli* may have for these alternative fusion protein substrates.

#### 3.4.6.1 Construction of (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-SUMO1-LmrC and (His)<sub>6</sub>-Ub-EGFP to investigate specificity of the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3)

The ubiquitin cDNA sequence was exchanged for the UBLs, NEDD8 or SUMO1. Changing only the UBL within the (His)<sub>6</sub>-Ub-LmrC fusion protein could potentially reveal any possible high affinity the prokaryotic DUB activity could have for either of the UBLs. The change from ubiquitin to NEDD8 would be a subtle one regarding size and 3D characteristics, as NEDD8 has a high sequence identity (58%) to ubiquitin, is identical in size to ubiquitin and carries the same features as ubiquitin, the GG C-terminal and ubiquitin superfold. A preference for ubiquitin over NEDD8 may indicate that ubiquitin is the true substrate for this prokaryotic DUB enzyme or enzymes with a function related to subverting host cell systems. The residue following NEDD8 and SUMO1 was S, just as it was in the fusion protein with ubiquitin and the residue following Ub in the (His)<sub>6</sub>-Ub-EGFP fusion protein was M.

The constructs encoding (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-SUMO1-LmrC and (His)<sub>6</sub>-Ub-EGFP were made as follows (detailed in 2.6.1). NEDD8 (256 bp), SUMO1 (322 bp) and (His)<sub>6</sub>-Ub-EGFP (990 bp) cDNAs were amplified to include appropriate restriction sites. The resulting DNA products and pHUE vector were separated by agarose gel electrophoresis, see Fig.3.10.A and double digested with the

appropriate enzymes (labelled 'Cut' in Fig.3.10.B and Fig.3.10.C). Gel purified inserts were then ligated with the appropriately cut vector which was subsequently used to transform *E. coli* substrain DH5 $\alpha$  with the ligation product. The purified plasmid from the resulting colonies was then screened for the presense of an insert using the appropriate restriction enzymes in a double digest (Fig.3.10.D). It was evident that all three constructs were coded for by the clones being screened, verified by sequencing. A very faint band was shown for double digested NEDD8 (227 bp) and SUMO1 (290 bp). A stronger band was seen for (His)<sub>6</sub>-Ub-EGFP (978 bp). Purified plasmid was then used to transform *E. coli* substrain Rosetta™2(DE3) for induction of recombinant protein expression.



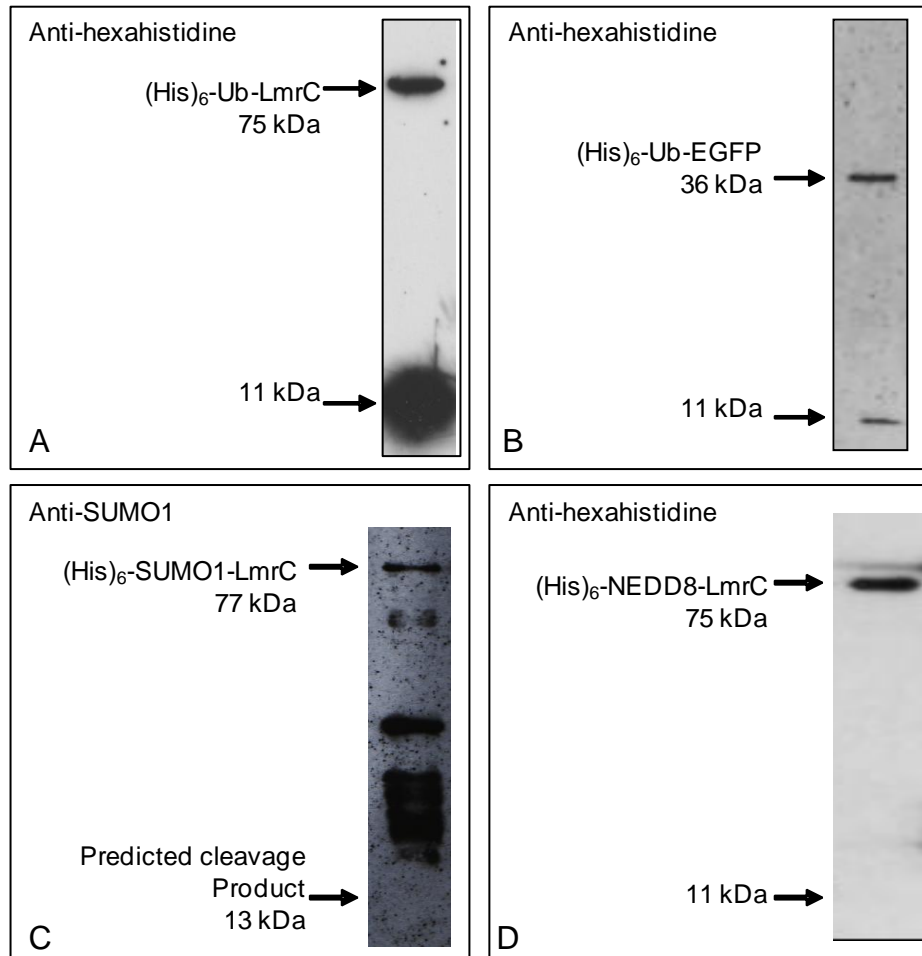
**Fig.3.10** All figures (A – D) show DNA electrophoresed on 1% (w/v) agarose gels. The three constructs generated for this study, (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-SUMO1-LmrC and (His)<sub>6</sub>-Ub-EGFP were generated first, by amplifying NEDD8, SUMO1 and (His)<sub>6</sub>-Ub-EGFP with restriction sites, shown in (A). The PCR products and vector, pHUE containing (His)<sub>6</sub>-Ub-LmrC, shown in (B and C) were then double digested (cut) or left uncut. Double digested DNA products were ligated and used to transform *E. coli* DH5 $\alpha$ , then grown over night on LB agar plates + ampicillin (100  $\mu$ g/ml) (37°C). Resulting colonies were grown over night in 10 ml cultures (37°C) and plasmid extracted using a kit, QIAprepSpin (QIAGEN). Successful clones were screened using a double digest to reveal the presence of the insert (size/bp) in the form of a second band (D). In this case the gels for NEDD8 and SUMO1 are the negative images in order to show the inserts which are faint bands.

3.4.6.2 Expression of fusion proteins, (His)<sub>6</sub>-SUMO1-LmrC, (His)<sub>6</sub>-NEDD8-LmrC and (His)<sub>6</sub>-Ub-EGFP in *E. coli* substrain Rosetta™2(DE3)

Recombinant proteins, including (His)<sub>6</sub>-Ub-LmrC were expressed in *E. coli* substrain Rosetta™2(DE3), purified and analysed by SDS-PAGE on a gradient gel, 5-20% (w/v) acrylamide and transferred on to nitrocellulose membrane and probed for either hexahistidine (Fig.3.11.A, B and D) or SUMO1 (Fig.3.11.C). The (His)<sub>6</sub>-Ub-EGFP and (His)<sub>6</sub>-Ub-LmrC fusion proteins were shown to have undergone cleavage to produce a single band (11 kDa). There was no band for the expected DUB cleavage product for either (His)<sub>6</sub>-NEDD8-LmrC (11 kDa) or (His)<sub>6</sub>-SUMO1-LmrC (13 kDa), however there were many bands noted below (His)<sub>6</sub>-SUMO1-LmrC which could be products of non-specific proteolysis. The cleavage of the (His)<sub>6</sub>-Ub-EGFP as well as WT (His)<sub>6</sub>-Ub-LmrC could imply that ubiquitin is very similar to the natural substrate for the *E. coli* DUB activity. The UBL, NEDD8 is very close to ubiquitin in sequence (58% identity) and is the same size (76 aa), but curiously NEDD8 was not cleaved from its fusion partner, LmrC. The lack of cleavage of the (His)<sub>6</sub>-SUMO1-LmrC fusion protein would suggest that the DUB activity in *E. coli* substrain Rosetta™2(DE3) does not normally target any protein analogous to SUMO1. These results provide evidence for a conservative DUB activity, expressed by *E. coli* substrain Rosetta™2(DE3) which has a natural substrate close to ubiquitin. The next thing considered was the possibility that this DUB



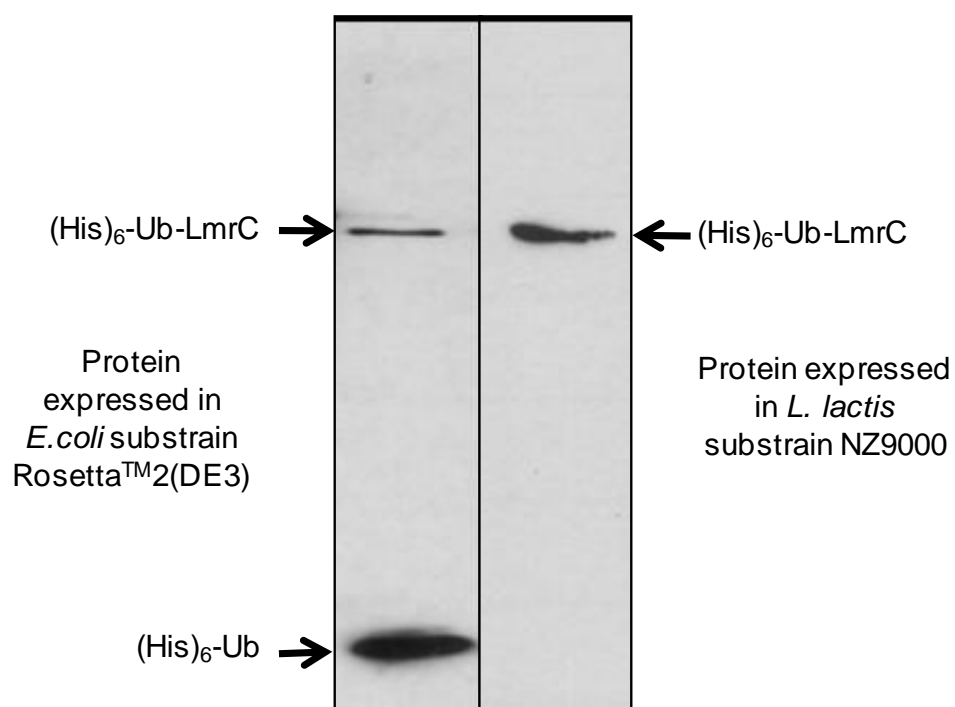
activity was only observed in *E. coli* and not in other bacterial strains; to test this supposition *Lactococcus lactis* was included in the study.



**Fig.3.11** *E. coli*, Rosetta™2(DE3) cells carrying a pHUE vector, sub-cloned with either, (His)<sub>6</sub>-Ub-LmrC, (His)<sub>6</sub>-Ub-EGFP, (His)<sub>6</sub>-SUMO1-LmrC or (His)<sub>6</sub>-NEDD8-LmrC were grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (50 µM for the NEDD8 and SUMO1 carrying clones and 100 µM for the EGFP clones) was used to induce protein expression. The resulting proteins were purified using cobalt affinity chromatography, electrophoresed on a 5-20% (w/v) gradient gel, transferred on to nitrocellulose membrane and probed for either hexahistidine (A, B and D) or SUMO1 (C). The (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP fusion proteins underwent DUB cleavage to produce a single band (11 kDa). There was no band for the expected cleavage products for either His)<sub>6</sub>-NEDD8-LmrC (11 kDa) or (His)<sub>6</sub>-SUMO1-LmrC (13 kDa). It is evident that there are many bands in (C) which could be products of non-specific proteolysis.

### **3.4.7 Investigating the possibility that DUB activity in *E. coli* was species specific**

The observed DUB activity could be exclusive to the *E. coli* bacterial strain. To investigate this possibility the *Lactococcus lactis* substrain, NZ9000 a gram positive, lactic acid bacterium was also analysed. Recombinant protein expression of (His)<sub>6</sub>-Ub-LmrC sub cloned into the pNZ8048 vector (constructed by J. Dorrian, School of Biomedical Sciences) was induced in this strain using nisin (2.1.2) and purified using cobalt affinity chromatography. Purified proteins were then analysed by SDS-PAGE on a gradient gel, 5-20% (w/v) acrylamide, and transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.3.12). The (His)<sub>6</sub>-Ub-LmrC fusion protein, when expressed in *L. Lactis* substrain NZ9000 was not cleaved at all (Fig.3.12.B) compared with when it was expressed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) (Fig.3.12.A). This indicates that the DUB activity is species specific.



**Fig.3.12** The WT fusion protein, (His)<sub>6</sub>-Ub-LmrC was expressed in one of two different bacterial strains, *E. coli* substrain Rosetta™2(DE3) (A) or *Lactococcus lactis* substrain NZ9000 (B). The resulting protein was then purified using cobalt affinity chromatography, electrophoresed on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and probed for ubiquitin. An immunoreactive band (11 kDa) was evident in proteins generated in *E. coli* (A) but not from those generated in *L. lactis*, suggesting that this DUB activity was strain specific.

### **3.5 Discussion**

This work shows that the *E. coli* substrain Rosetta™2(DE3) has an intrinsic, specific DUB activity similar to eukaryotic DUBs. The cleavage of (His)<sub>6</sub>-Ub-LmrC fusion protein by the DE3 strain of *E. coli* observed by other members of this group (H. Jeffery and I. Kerr) was shown to be reproducible in the present study when expressed in *E. coli* substrain Rosetta™2(DE3) (Fig.3.2). The precision of this DUB cleavage was confirmed using N-terminal sequence analysis (Fig.3.3) and MS (Fig.3.7, Table 3.1). These findings agree with the observations of other workers and with Catanzariti *et al.* who also used N-terminal sequencing analysis to confirm specificity of intrinsic DUB cleavage in *E. coli* substrain AR58. In the present study, the use of (His)<sub>6</sub>-Ub-LmrC mutants, GA, GP and GGP, mutated at the DUB site to investigate specificity, ultimately showed the prokaryotic DUB activity had a 'DUB activity similar to that of eukaryotic DUBs. However, there is a caveat associated with the use of this particular fusion protein. In addition to the DUB site within (His)<sub>6</sub>-Ub-LmrC we showed that there was also a second, non-specific cleavage site 10 aa, C-terminal to the DUB site in LmrC (see cartoon in Fig.3.8).

A cleavage product, slightly higher in mass than the (His)<sub>6</sub>-Ub (~11 kDa) was observed in the GA, GP and GGP samples analysed for MS MALDI-TOF in Fig.3.7. The resulting masses for proteins found in the

11 kDa region of all the samples analysed were compared with estimated masses for possible DUB cleavage products (Table 3.1). This clearly demonstrated that the bands observed in Fig.3.5, lanes indicated GA, GP and GGP, were not the products of DUB cleavage. Instead, these were products from a non-specific cleavage site 10 aa C-terminal to the DUB site. It is interesting to note that there is no evidence in the data obtained from either N-terminal sequencing analysis or MS that WT (His)<sub>6</sub>-Ub-LmrC undergoes cleavage at the second site. This absence of cleavage of the second site within WT (His)<sub>6</sub>-Ub-LmrC could imply that for some reason the second site is only accessible while the DUB site is not being processed. It is likely that when the DUB site is not accessible that this drives the use of the second non-specific site by another bacterial enzyme. It is also possible that the second cleavage site is rendered unobtainable to the non-specific enzyme due to steric hindrance from the DUB activity as the two sites are only 10 aa apart. It is also possible that the mechanism of proteolysis that takes place depends upon some signal for example the removal of an N-terminal residue by the N-end rule. For example the DUB activity may be favoured if it occurred co-translationally while the non-specific protease may have to have an additional signal for proteolysis in the form of the removal of the N-terminal M by the N-end rule.

The finding in the present study that the GG-P bond was not cleaved by the DUB activity in *E. coli* substrain, Rosetta™2(DE3) implies that this prokaryotic DUB activity has an activity comparable to eukaryotic DUBs. The lack of cleavage of the GP mutant fusion protein by the DUB activity in Rosetta™2(DE3) shows that the catalytic mechanism is similar to that of a mammalian DUB (D. A. Gray personal communication). Gray and co-workers expressed a Ub<sup>G76P</sup> mutant within the fusion protein, (His)<sub>6</sub>-Ub<sup>G76P</sup>-EGFP in mammalian cells and found it also resisted cleavage. Interestingly, in the present study, GA resisted cleavage in *E. coli* substrain Rosetta™2(DE3) which is a characteristic more comparable with yeast DUBs. T. R. Butt reported the resistance of a fusion protein, Ub<sup>G76A</sup>-metallothionein to cleave when expressed in yeast<sup>148</sup>. The (His)<sub>6</sub>-Ub<sup>G76A</sup>-EGFP fusion protein offered no such challenge to the mammalian cell system (D. A. Gray) further highlighting differences between prokaryotic and eukaryotic cells.

Characterisation of the DUB activity in *E. coli* substrain Rosetta™2(DE3) was investigated by replacing either one of the fusion proteins within the (His)<sub>6</sub>-Ub-LmrC fusion protein. The DUB activity in *E. coli* substrain Rosetta™2(DE3) did cleave (His)<sub>6</sub>-Ub-EGFP but not (His)<sub>6</sub>-NEDD8-LmrC or (His)<sub>6</sub>-SUMO1-LmrC (Fig.3.11). The amino acid which follows immediately after G76 of the UBL at the DUB site within

(His)<sub>6</sub>-NEDD8-LmrC and (His)<sub>6</sub>-SUMO1-LmrC is S, the same as in WT (His)<sub>6</sub>-Ub-LmrC. The amino acid following the last G in the UBL at the DUB site within (His)<sub>6</sub>-Ub-EGFP was M. The ability of the DUB activity to cleave (His)<sub>6</sub>-Ub-EGFP as well as WT (His)<sub>6</sub>-Ub-LmrC would appear to denote a close similarity between ubiquitin and the natural substrate for this prokaryotic DUB activity. Nonetheless, neither of these fusion proteins underwent complete DUB processing, as was evident by the presence of the intact fusion proteins in both cases (Fig.3.11 & Fig.3.5 A&C). The absence of cleavage of either of the fusion proteins containing UBLs, NEDD8 or SUMO1, (His)<sub>6</sub>-NEDD8-LmrC or (His)<sub>6</sub>-SUMO1-LmrC respectively, could also signify a close relationship of ubiquitin to the natural substrate for the prokaryotic DUB activity. There is a high sequence identity when comparing NEDD8 with ubiquitin (58% identity) and both UBLs have the features associated with ubiquitin, a flexible diglycine C-terminus and the same three-dimensional structure with a  $\beta$ -grasp fold.

A constraint over the complete processing of the fusion proteins could be due to a lack of specificity of the prokaryotic DUB activity for a particular 'sub-type' of a peptide bond. During translation of ubiquitin gene products (precursors) the  $\alpha$ -carboxyl group of ubiquitin becomes condensed with the  $\alpha$ -amino group of the neighbouring ubiquitin. Conversely when ubiquitin is attached post translationally to a protein,



the  $\alpha$ -carboxyl group of ubiquitin is condensed the  $\varepsilon$ -amino group of the target protein, forming an isopeptide bond. Some DUBs discriminate between the two types of bonds and it is possible that the DUB activity observed in *E.coli* substrain Rosetta™2(DE3) prefers isopeptide bonds. Of course it is also possible that more than one type of DUB activity occurs in *E. coli*. Another possible restriction over the complete processing of (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP fusion proteins by the DUB activity could be explained by the ability of LmrC and EGFP to form homo-dimers. It is possible that the cleaved LmrC or EGFP can bind to intact fusion protein, preventing access by the DUB to the DUB site. In Fig.3.5 (A) it is clear that there is a near equal amount of intact and cut (His)<sub>6</sub>-Ub-LmrC protein and that this equal stoichiometry is evidence that of such a dimer.

This study has revealed that the DUB activity observed in *E. coli* substrain Rosetta™2(DE3) was strain specific. No DUB processing of WT (His)<sub>6</sub>-Ub-LmrC occurred when this protein was expressed in *L. lactis* (Fig.3.12). The second site was also not cleaved in *L. lactis* which shows a real difference between the strains of bacteria. However, the possibility that the DUB activity is also substrain specific has not been ruled out as DUB activity has only been observed in or associated with, the *E. coli* substrains; Rosetta™2(DE3) in the current study and in DH5 $\alpha$ , AR58 and BL21 during work carried out by other groups<sup>3, 4, 138</sup>.

This work has demonstrated that the *E. coli* substrain, Rosetta™2(DE3) expressed a DUB or DUBs with a specific DUB mechanism similar in some ways to conventional eukaryotic DUBs, in not cleaving GG-P bonds. However, this bacterial DUB activity shares characteristics in its catalytic mechanism with both mammalian and yeast DUBs as it is unable to cleave either Ub<sup>G76P</sup> or Ub<sup>G76A</sup>, respectively when expressed within a fusion protein. Further work needs to be carried out to determine if the bacterial DUB activity shows discrimination between the two types of bond, peptide, or isopeptide. Additionally, a candidate approach may allow identification of the DUB enzyme or enzymes responsible for the DUB activity observed.

## **4 A candidate approach to identify the prokaryotic enzyme responsible for the DUB activity in *E. coli* substrain Rosetta™2(DE3)**

### **4.1 Introduction**

In Chapter 3, it was first established that the DUB activity in a DE3 strain of *E. coli* was reproducible in *E. coli* substrain Rosetta™2(DE3) and occurred using specific DUB cleavage. The DUB activity was then shown to have an activity similar to eukaryotic DUBs, comparable to eukaryotic DUBs in addition to preferring ubiquitin as a substrate over SUMO1 and NEDD8. This having been achieved, the impetus was then to establish the identity of the DUB activity in *E. coli* substrain Rosetta™2(DE3) using a candidate approach.

The first DUB candidate considered was the *E. coli* elongation factor EF-Tu. This was a natural starting point following on from a report by A. Ciechanover and colleagues<sup>139</sup>. Their work concerned the characterisation of a protein, factor Hedva (FH) in which they also reported that the eukaryotic elongation factor, EF-1 $\alpha$  had a DUB activity, *in vitro* which deubiquitinated a histone, H2A. Additionally, the authors noted that the prokaryotic elongation factor, EF-Tu, expressed by *E. coli* substrain DH5 $\alpha$ , was also associated with this DUB activity,

as EF-Tu co-migrated with H2A. Therefore this DUB activity was associated with cleavage of an isopeptide bond rather than with the peptide bonds so far considered in the current study. In the present study, it was shown in Chapter 3 that not all of the substrate was cleaved by the DUB activity in *E. coli* substrain Rosetta™2(DE3). This lead to the speculation that the DUB activity in *E. coli* substrain Rosetta™2(DE3) may prefer an isopeptide bond rather than a peptide bond found in linear fusion proteins. The association of EF-Tu with a DUB activity made this a good candidate to investigate in the identification of the DUB activity observed in our *E. coli* substrain Rosetta™2(DE3) study. Eukaryotic (rat) EF-1 $\alpha$  Type 2 (EF1 $\alpha$ 2) was also included in this study, to compare both eukaryotic and prokaryotic elongation factors.

Whilst this work was in progress the second DUB candidate considered was the *E. coli* protein, *elaD*. A group led by H. L. Ploegh, discovered a protein, *elaD* expressed in *E. coli* substrain K12, BL21 with DUB activity<sup>14</sup>. Bioinformatics was used to search for new members of the clan containing ubiquitin and ubiquitin-like proteases (clan CE) in viruses, bacteria and eukaryotes, which revealed the protein, *elaD* as a potential protease. Two subsequent biochemical tests, carried out *in vitro*, showed that GST-*elaD* displayed DUB activity with substrates termed suicide inhibitors as well as with the fluorogenic substrate Ub-

AMC. The suicide inhibitors used were mouse UBLs, ubiquitin and SUMO-1 and Nedd8 which were modified to have a 'diverting' group for attack, a Michael acceptor, adjacent to the nucleophilic centre normally favoured by DUBs. The use of suicide inhibitors demonstrated that elad favoured ubiquitin as a substrate over SUMO-1 and Nedd8. This particular preference for Ub as a substrate is a characteristic that elad shares with the DUB activity investigated in the current study. As a result of these findings elad was included as a candidate for the DUB activity in *E. coli* substrain Rosetta™2(DE3).

The work covered in this chapter discusses cloning of the two DUB candidates, EF-Tu and elad, a gene knock-out of elad and a ubiquitin-Sepharose (ubiquitin-Sepharose) binding assay to select for candidates that bind to mammalian ubiquitin.

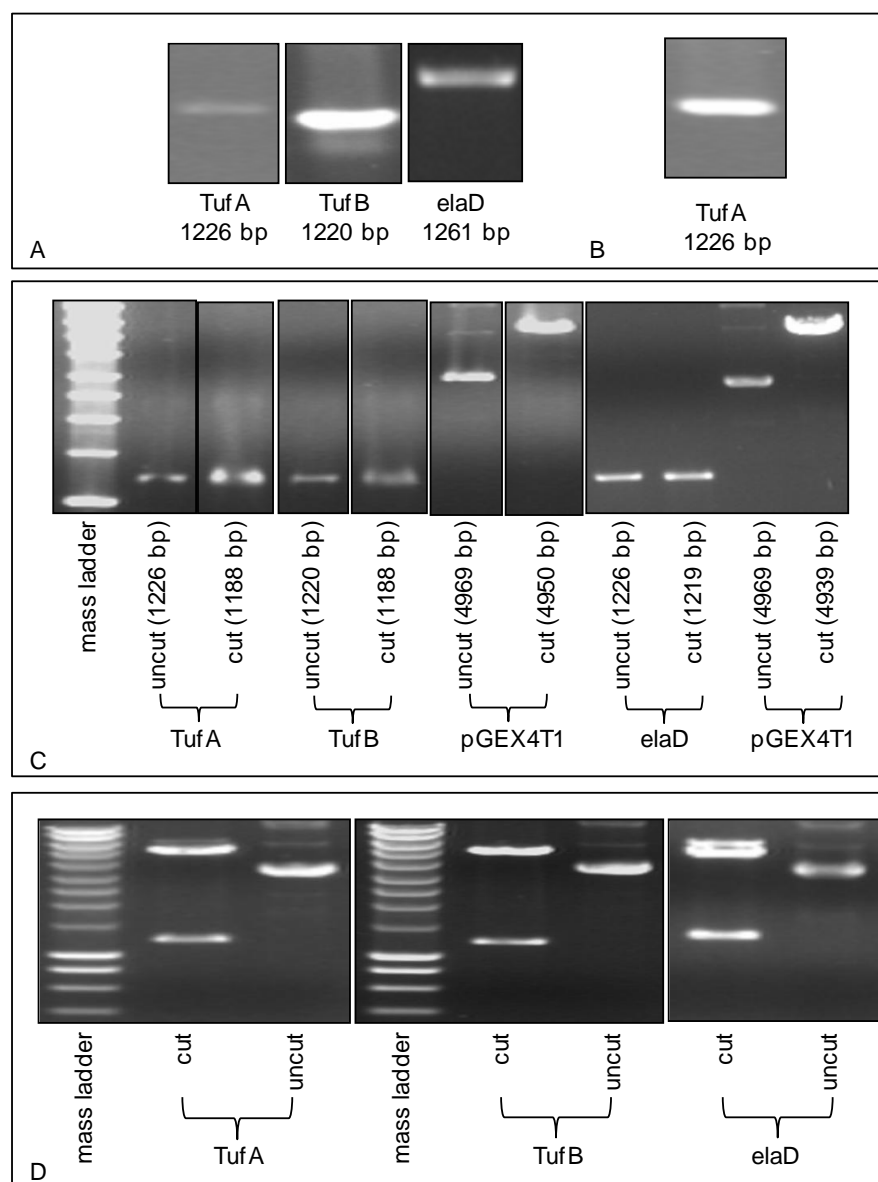
## **4.2 Results**

### **4.2.1 The cloning of GST-tagged candidates EF-Tu and elaD, for the DUB activity in *E. coli* substrain Rosetta™2(DE3)**

The aim of the work discussed in this chapter was to use a candidate approach in an attempt to identify the DUB activity in *E. coli* substrain Rosetta™2(DE3) observed in Chapter 3. The initial objective was to create constructs for the over-expression of GST-tagged DUB candidates, EF-Tu and elaD in *E. coli* substrain XL10Gold. The next objective was to use the GST-tagged candidates in ubiquitin binding assays *in vitro*, to see if any candidate binds to ubiquitin (many DUBs have this property).

Initially, the DNA sequence for elaD in *E. coli* substrain Rosetta™2(DE3) was sequenced then used in an alignment with the sequence for elaD (in *E. coli* substrain K12) reported by A. Catic<sup>14</sup>. This confirmed that the elaD sequences from Rosetta™2(DE3) and K12 were identical. The cDNAs encoding EF-Tu (43 kDa) and elaD (45 kDa) proteins were cloned into pGEX4T1 in three main steps, detailed in 2.6.2. In the first step, inserts were amplified by PCR. Two genes code for EF-Tu in *E. coli* '*TufA* and *TufB*' and the proteins translated from them, TufA and TufB differ only in the last amino acid residue: TufA has G and TufB has S. The amplified *TufA* and *TufB* inserts are shown in

Fig.4.1.A and amplified *elaD* in 4.1.B. The resulting *TufA* and *TufB* PCR products were then used as a DNA template for a subsequent PCR Fig.4.1.B to increase product concentration. In the second step, purified PCR products and vectors were double digested using appropriate restriction enzymes and separated by electrophoresis Fig.4.1.C. In the third step, the purified, double digested DNA inserts and corresponding double digest vectors, were ligated and used to transform *E. coli* substrain XL10Gold. The purified vector from the resulting colonies was then screened using appropriate pairs of restriction enzymes in a digest. All three constructs coded for a candidate shown in Fig.4.1.D and were verified by DNA sequencing. The cDNAs encoding eukaryotic elongation factors, GST-EF-1 $\alpha$ 1 and GST-EF-1 $\alpha$ 2 proteins were cloned into pGEX4T1 by J. R. Cavey, University of Nottingham.



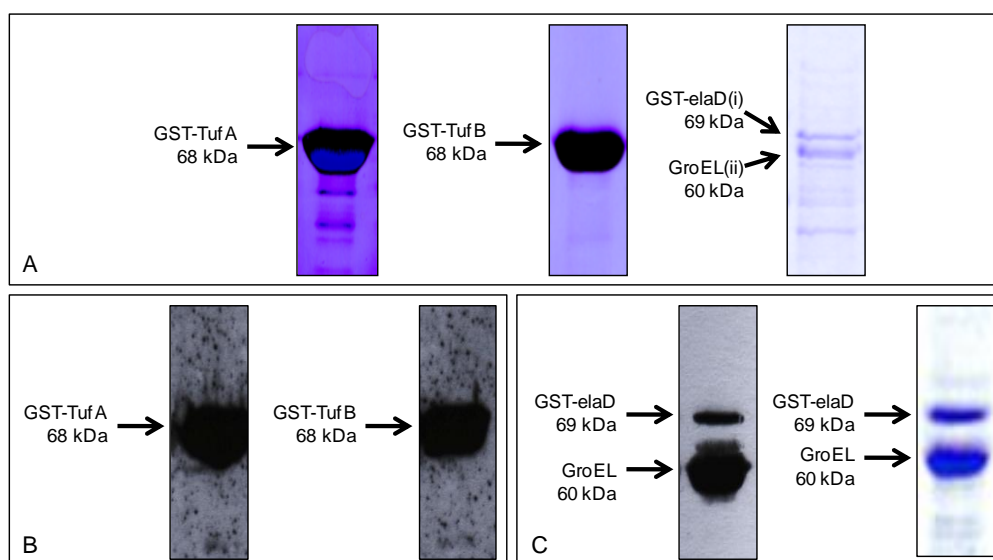
**Fig. 4.1** cDNAs encoding the candidates for the DUB activity, EF-Tu and elaD were inserted in to the pGEX4T1 vector to place a glutathione transferase (GST) tag at the N-terminal of the candidate. EF-Tu is coded for by two near identical coding regions, termed *TufA* and *TufB*. The *TufA* and *TufB* proteins differ only in the final amino acid, glycine in *TufA* and serine in *TufB*. (A – D) show DNA electrophoresed on 1% (w/v) agarose gels. First, candidates were amplified with restriction sites (A), *TufA* PCR product was used as a template in a subsequent PCR to increase DNA concentration (B). The PCR products and vector were then double digested 'Cut' or left uncut (C) with sizes (bp) indicated. Double digested DNA products were ligated and used to transform *E. coli* substrain XL10Gold, then grown over night (37°C) on LB agar plates (100 µg/ml ampicillin). Resulting colonies were grown over night in 10 ml cultures, with agitation (37°C) and plasmid extracted using a kit, Q1AprepSpin (QIAGEN). Successful clones were screened using a double digest to reveal the presence of the insert (of the same sizes indicated in C) in the form of a second band, see (D).



#### **4.2.2 Over-expression of DUB Candidate proteins: GST-TufA, GST-TufB, GST-elaD and eukaryotic elongation factors, GST-EF-1 $\alpha$ 1, GST-EF1- $\alpha$ 2**

GST-tagged proteins were over-expressed (2.1) in *E. coli* substrain XL10Gold. Purified GST-tagged protein (2.4.2.1) was eluted in gel loading buffer and electrophoresed on a gradient gel, 5-20% (w/v) acrylamide gradient (2.4.4) and either stained with Coomassie blue stain (Fig.4.2.A and Fig.4.2.C.2) or transferred on to nitrocellulose membrane and probed for GST (Fig.4.2.B and Fig.4.2.C.1). GST-TufA and GST-TufB bands (68kDa) are shown on the Coomassie stained gel (Fig.4.2.A) and when probed for GST (Fig.4.2.B). Two weak bands (~68 kDa) were seen on Coomassie stained gel (A), either could potentially be GST-elaD. Therefore the bands (i) and (ii), indicated in (Fig.4.2.B), were submitted for Peptide Mass Fingerprint analysis and resulting sequence matches showed that (i) represented GST-elaD and (ii) was a 60 kDa chaperone protein, GroEL (Cpn60). It is interesting that GroEL did not co-purify with TufA or TufB proteins. The proteins in (Fig.4.2.A) and (Fig.4.2.B) were expressed from a 10 ml culture and the weak band for GST-elaD shows that this is expressed at low levels compared with GST-TufA and TufB. The product from a 50 ml culture was then electrophoresed as described above resulting in a stronger band (~69 kDa) for GST-elaD on the Coomassie stained gel, which was also immunoreactive with the GST probe (Fig.4.2.C). There was a high level

of expression of GST-TufA and GST-TufB. This indicates a need to optimise GST-elaD expression and purification.



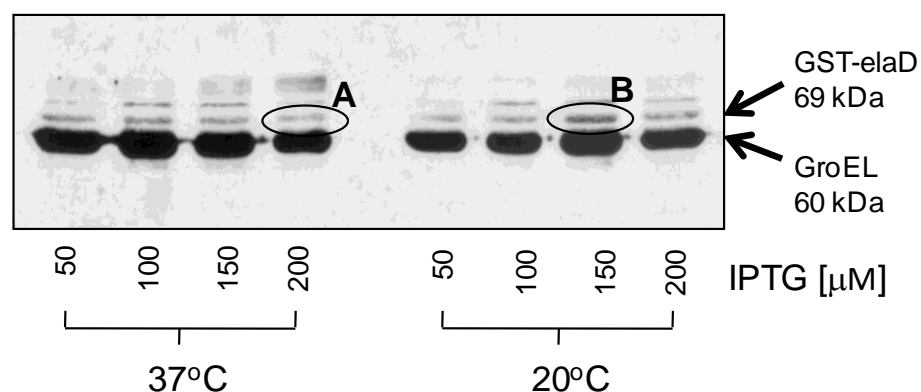
**Fig. 4.2** DUB candidate proteins, GST-TufA, GST-TufB and GST-elaD were expressed in *E. coli* substrain XL10Gold and grown in LB plus ampicillin (100  $\mu$ g/ml) until optical density = 0.6-0.8. IPTG (200  $\mu$ M) was used to induce protein expression. The resulting protein was purified using glutathione-Sepharose 4B beads, eluted in gel loading buffer, electrophoresed on a gradient gel 5-20% (w/v) acrylamide and either stained using Coomassie Blue stain (A and C.2) or transferred on to nitrocellulose membrane and probed for GST (B and C.1). Bands (68 kDa) representing GST-TufA and GST-TufB are shown on the Coomassie stained gel (A) as well as when probed for GST (B). Two weak bands (~69 kDa) were seen on Coomassie stained gel (A), one of which was thought to be GST-elaD. The bands indicated in (B), (i) and (ii) were submitted for Peptide Mass Fingerprint analysis and resulting sequences and subsequent sequence alignments showed (i) represented GST-elaD and (ii) was chain A of GroEL. The proteins in (A) and (B) were expressed from a 10 ml culture and the weak band for GST-elaD shows that this is expressed at low levels compared with GST-TufA and TufB. The recombinant protein from a 50 ml culture of XL10Gold was then purified and electrophoresed as described above resulting in a stronger band (~69 kDa) for GST-elaD on the Coomassie stained gel, which was also immunoreactive with the GST probe (C). It is evident that the antibody to GST cross-reacted with the GroEL protein (C).

### **4.2.3 Improving purification and yield of the DUB candidate, GST-elaD protein**

It was evident, in Fig.4.2 (A and C) that GST-elaD protein expression was relatively low compared with that of GST-TufA and GST-TufB, thus it was necessary to increase elaD protein yield. An effort was also made to remove the 60 kDa protein, GroEL which co-purified with GST-elaD.

#### **4.2.3.1 Increasing the induced yield of DUB candidate, GST-elaD protein**

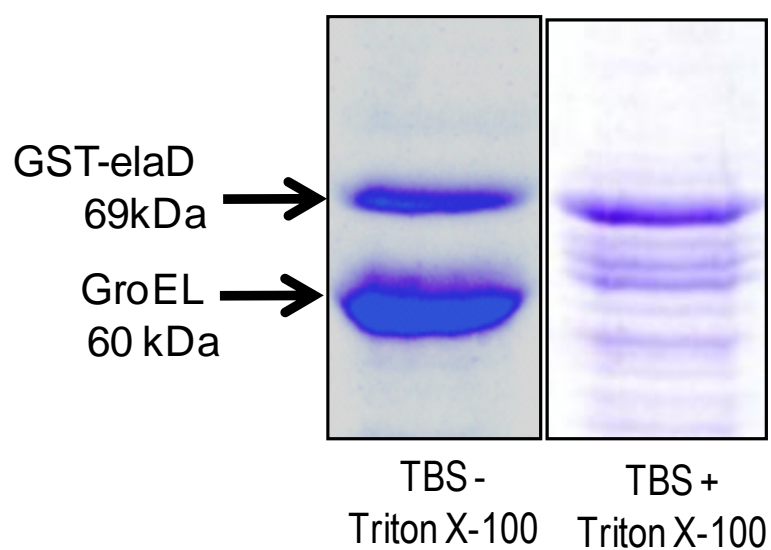
GST-elaD was expressed in *E. coli* substrain XL10Gold, grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8 (Fig. 4.3). A range of IPTG concentrations (50 µM - 200 µM) were used to induce protein expression at two different temperatures and times, 37°C, 3 hr or 20°C, over night. The resulting protein was purified using glutathione-Sepharose 4B beads, washed in TBS, eluted in gel loading buffer, electrophoresed on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and probed for GST. There was a limited improvement in yield of GST-elaD protein when the culture was induced with 150 µM IPTG, over night at 20°C (Fig. 4.3 band B), compared with conditions used previously, 200 µM IPTG, 3 hr at 37°C (Fig. 4.3 band A). Therefore the new growth conditions were chosen for future use.



**Fig. 4.3** To increase the yield of GST-elaD, protein expression was optimised by varying the concentration of IPTG and temperature. elaD protein was expressed in *E. coli* substrain XL10Gold, grown in LB plus ampicillin (100 μg/ml) until optical density = 0.6-0.8. A range of IPTG concentrations (50 μM - 200 μM) were used to induce protein expression at two different temperatures, 37°C, 3 hr or 20°C, over night. The resulting protein was purified using glutathione-Sepharose 4B beads, washed in TBS, eluted in gel loading buffer, electrophoresed on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and probed for GST. None of the conditions produced a significant increase in yield of elaD. However when comparing the band for maximal elaD protein (B) with the band of elaD induced in the conditions used previously, 200 μM IPTG, 37°C 3hr (A) there was a limited improvement in the induction of elaD protein. Therefore future inductions of GST-elaD were carried out at 150 μM IPTG, 20°C over night. The antibody to GST was also seen to cross-react with GroEL as it did in Fig.4.2.

#### 4.2.3.2 Removing GroEL from purified GST-elaD protein with the aim to produce homogenous pure protein

It was necessary to produce purified GST-elaD without GroEL as GroEL was present in the purified sample in relatively large amounts and the antibody to GST cross-reacts with GroEL. GST-elaD protein was expressed in *E. coli* substrain XL10Gold, grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (200 µM) was used to induce protein expression at 37°C for 3 hr. For protein purification, lysates were made up either as described in (2.1.3) or with 0.1% (v/v) Triton X-100 in TBS. Other washes (High salt, a mild non-ionic detergent: bacterial protein extraction reagent (BPER) (ThermoScientific) and ATP) were also tested, data not included (2.4.2.1) but they were not conducive to protein-protein interaction studies or enzyme studies. The resulting purified proteins were electrophoresed on a gradient gel 5-20% (w/v) acrylamide and stained with Coomassie blue stain (Fig. 4.4). Triton X-100 was sufficient in removing GroEL protein from the purified GST-elaD as can be seen when compared with TBS - Triton X-100 and therefore was chosen for subsequent elaD protein purification.



**Fig. 4.4** Protein purification conditions were varied to remove GroEL, a 60 kDa protein which co-migrated with GST-elaD. GST-elaD protein was expressed in *E. coli* substrain XL10Gold, grown in LB plus ampicillin (100  $\mu$ g/ml) until optical density = 0.6-0.8. IPTG (200  $\mu$ M) was used to induce protein expression, 37°C, 3 hr. The resulting protein was purified using glutathione-Sepharose 4B beads in TBS, either in the presence, or absence of 0.1% (v/v) Triton X-100. Purified GST-elaD was electrophoresed on a gradient gel 5-20% (w/v) acrylamide and stained using Coomassie Blue stain. Triton X-100 was used in subsequent purifications of GST-elaD as it succeeded in washing off GroEL.

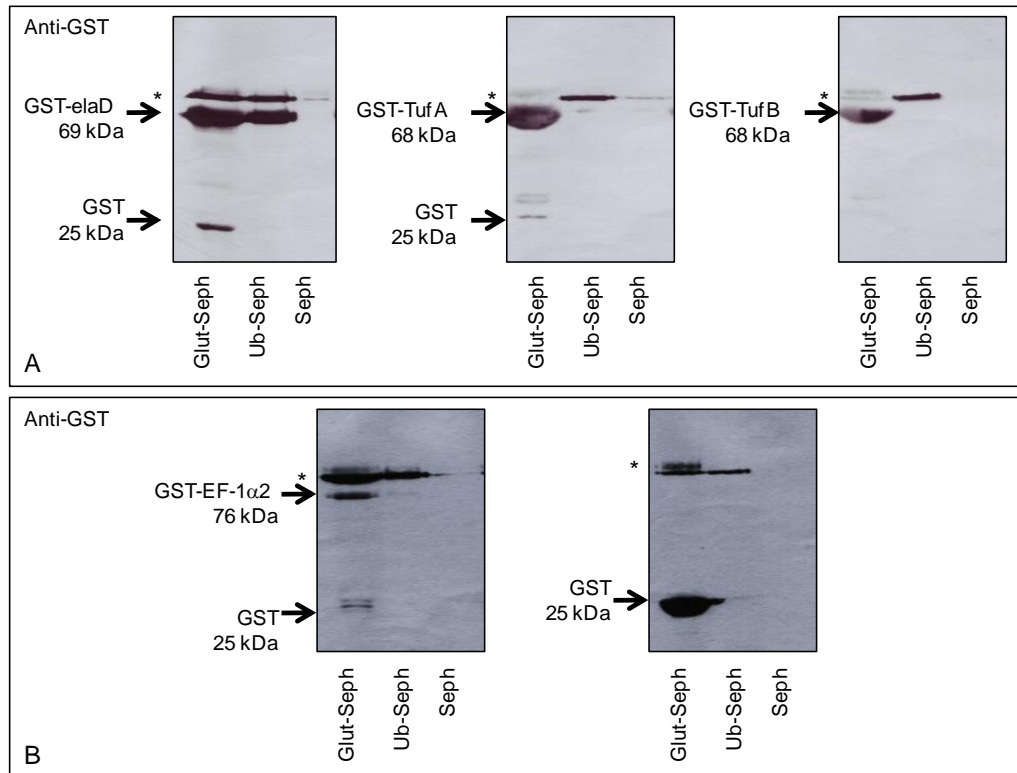
#### **4.2.4 An assay to screen the DUB candidates for the ability to bind with ubiquitin**

In Chapter 3, an *in vivo* DUB activity was observed when (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP substrates were expressed in Rosetta<sup>TM</sup> 2(DE3). This meant that the DUB activity could recognise and bind to ubiquitin in some way in order to cleave the ubiquitin fusion proteins. Therefore an initial assay was chosen to investigate and triage DUB candidates, TufA, TufB and elaD using a test for the ability to bind to ubiquitin in a pull-down assay. As well as the DUB candidates, eukaryotic EF-1 $\alpha$ 2 was also included in this study, to compare both eukaryotic and prokaryotic elongation factors and GST as a control.

GST-tagged proteins were over-expressed as described in 2.1, with the exception that GST-elaD induction was carried out using 150  $\mu$ M IPTG followed and incubated (20°C) for 15 hr. Lysates were prepared in 0.1% (v/v) Triton X-100 in TBS as described 2.1.2. Protein was captured using a bead slurry (100  $\mu$ l) of either, glutathione-Sepharose 4B beads (glutathione-Sepharose), ubiquitin-Sepharose or Sepharose and incubated (4°C) rotating for 1 hr, then washed with 0.1% (v/v) Triton X-100 in TBS. Proteins were eluted off the beads using gel loading buffer, electrophoresed on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and probed for GST (Fig.4.5). A non-specific protein indicated (\*) which cross reacted with the GST antibody,



was present for all candidates as well as EF-1 $\alpha$ 2 and GST. All candidates were captured on glutathione-Sepharose 4B beads and it is evident that GST-elaD was the only DUB candidate to bind to ubiquitin-Sepharose. None of the candidates, GST or EF-1 $\alpha$ 2 bound to Sepharose control beads. These results provide evidence for elaD being a candidate for the DUB activity observed in Rosetta<sup>TM</sup>2(DE3). Therefore the next step was to create an elaD gene knock-out in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) and to express (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP in Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>apr</sup> to determine if the *in vivo* DUB activity ceased in the absence of elaD expression.

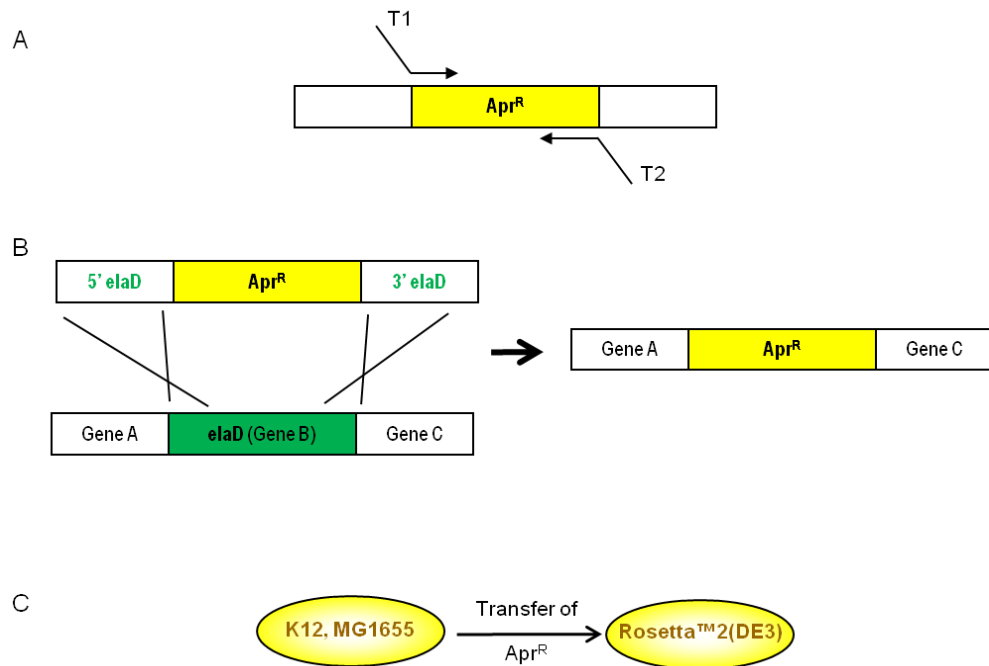


**Fig. 4.5** *E. coli* substrain XL10Gold, transformed with constructs coding for GST-tagged DUB candidates or GST as indicated were grown in LB plus ampicillin (100 µg/ml) until optical density = 0.6-0.8. IPTG (150 µM) was used to induce GST-elaD protein expression, 200 µM IPTG was used to induce the other GST-tagged DUB candidates and GST. Cultures expressing GST-elaD were incubated (20°C) for 15 hr, other DUB candidate and GST expressing cultures were incubated (37°C) for 3 hr. Lysates were prepared in 0.1% (v/v) Triton X-100 in TBS as described 2.1.2. Protein was captured using either glutathione-Sepharose 4B beads (Glut-Seph), Ub-Sepharose or Sepharose control beads and incubated (4°C) rotating for 1 hr, then washed with 0.1% (v/v) Triton X-100 in TBS. Proteins were eluted off the beads using gel loading buffer, electrophoresed on a gradient gel 5-20% (w/v) acrylamide, transferred on to nitrocellulose membrane and probed for GST. A non-specific protein indicated (\*) was present for all candidates including EF-1α2. All candidates were captured on glutathione-Sepharose beads and GST-elaD was the only DUB candidate to bind to ubiquitin-Sepharose. None of the DUB candidates, GST or EF1α2 bound to Sepharose control beads. This provides evidence for elaD being a strong candidate for the DUB activity observed previously in Rosetta<sup>TM</sup>2(DE3).

#### **4.2.2 Generation of the *elaD* knock-out, Rosetta<sup>TM</sup> 2(DE3) $\Delta$ *elaD*<sup>Apr</sup> to investigate further the possibility that *elaD* may be responsible for the DUB activity observed in this *E. coli* substrain**

In order to generate the *elaD* gene knock-out in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) a three step method was used. This method involved knocking out *elaD* in one strain of *E. coli* (K12) and then this gene knock-out characteristic was transferred to the desired strain of *E. coli* (Rosetta<sup>TM</sup>2(DE3)), developed by K. A. Datsenko and B. L. Wanner and subsequently adapted by E. Bolt (Fig. 4.6)<sup>151</sup>.

First the antibiotic cassette, in this case coding for resistance to apramycin (*Apr*<sup>R</sup>) was amplified using PCR to include the beginning and the end of *elaD*, the sequence to be knocked out (Fig.4.6.A). Within the second step, the *E. coli* substrain K12, MG1655 (JIG182) underwent two subsequent transformations, first with pKD46 then with *Apr*<sup>R</sup> DNA. The vector, pKD46 codes for the  $\lambda$  Red enzymes which carry out homologous recombination, essentially replacing DNA sequences, *elaD* with an *Apr*<sup>R</sup> cassette (Fig.4.6.B) to generate JIG182 $\Delta$ *elaD*<sup>Apr</sup>. Colonies were screened for JIG182 $\Delta$ *elaD*<sup>Apr</sup> using colony PCR. In the final step, Phage 1 transduction (also involving homologous recombination) was used to transfer the *elaD* gene knock-out characteristic from JIG182 $\Delta$ *elaD*<sup>Apr</sup> to create Rosetta<sup>TM</sup> 2(DE3) $\Delta$ *elaD*<sup>Apr</sup>. Colony PCR was used to screen for Rosetta<sup>TM</sup> 2(DE3) $\Delta$ *elaD*<sup>Apr</sup>.



**Fig. 4.6** There were three main steps, in knocking the *elaD* gene out from *E. coli* substrain Rosetta™2(DE3), which are briefly outlined above (A-C). The first step involved amplification of the apramycin resistance cassette (*Apr<sup>R</sup>*), using PCR to include the 5' and 3' regions of the *elaD* gene to be knocked out (A). In the second step, homologous recombination was used, to replace *elaD* with an *Apr<sup>R</sup>* cassette, a process carried out by  $\lambda$  Red recombinase enzymes coded for by the PKD46 vector (B). This process involved carrying out two subsequent transformations of *E. coli* substrain, K12 MG1655, first using PKD46 vector DNA and secondly, using PCR product from (A). The resulting colonies were screened using colony PCR. In the final step (C), phage 1 transduction was used to convey the *Apr<sup>R</sup>* characteristic in K12 to Rosetta™2(DE3). This was done by incubating phage 1 viral lysate with the K12 carrying *Apr<sup>R</sup>*, centrifuging (10 000 rpm) 20 min, incubating the resulting supernatant with Rosetta 2 (DE3) and finally screening for the *elaD* knock out using colony PCR.

Fig.4.7 shows the results from each of the 3 main steps used to ultimately develop Rosetta™ 2(DE3)ΔelaD<sup>Apr</sup>. In step 1 (Fig.4.7), PCR was used to amplify the antibiotic resistance cassette, in this case the Apr<sup>R</sup> sequence, using primers which also included regions homologous to the first and last 30 nucleotides of the *elaD* gene targeted for the knock-out. Two subsequent PCRs were carried out, PCR1 and PCR2, using the same conditions except PCR1 product was used as a template in PCR2 to increase DNA concentration. PCR primers, 'ElaDKOF' and 'ElaDKOR' were used to amplify the Apr<sup>R</sup> sequence using Pfu DNA polymerase (2.2.4.1). Template DNA for PCR1 was prepared using *E. coli* substrain K12, MG1655 (JIG-515) carrying Apr<sup>R</sup> sequence, as described in 2.2.3.1. There are strong bands for Apr<sup>R</sup> DNA amplified in PCR2, compared with the weaker Apr<sup>R</sup> DNA band amplified in PCR1. The bands from PCR2 were then DpnI digested (2.2.5.2) to remove the template and therefore any chance of transferring the Apr<sup>R</sup> in the absence of generating JIG182ΔelaD<sup>Apr</sup>. DpnI digested DNA was gel purified (2.2.5.1) and concentrated via ethanol precipitation (2.2.8) to give 2.6 μg of gel purified amplified Apr<sup>R</sup> DNA which was resuspended in 7 μl sterile distilled water ready for step 2.

In Step 2, *elaD* was knocked out in the *E. coli* substrain K12, MG1655 (JIG-182) creating JIG182ΔelaD<sup>Apr</sup>. To achieve this JIG-182 underwent

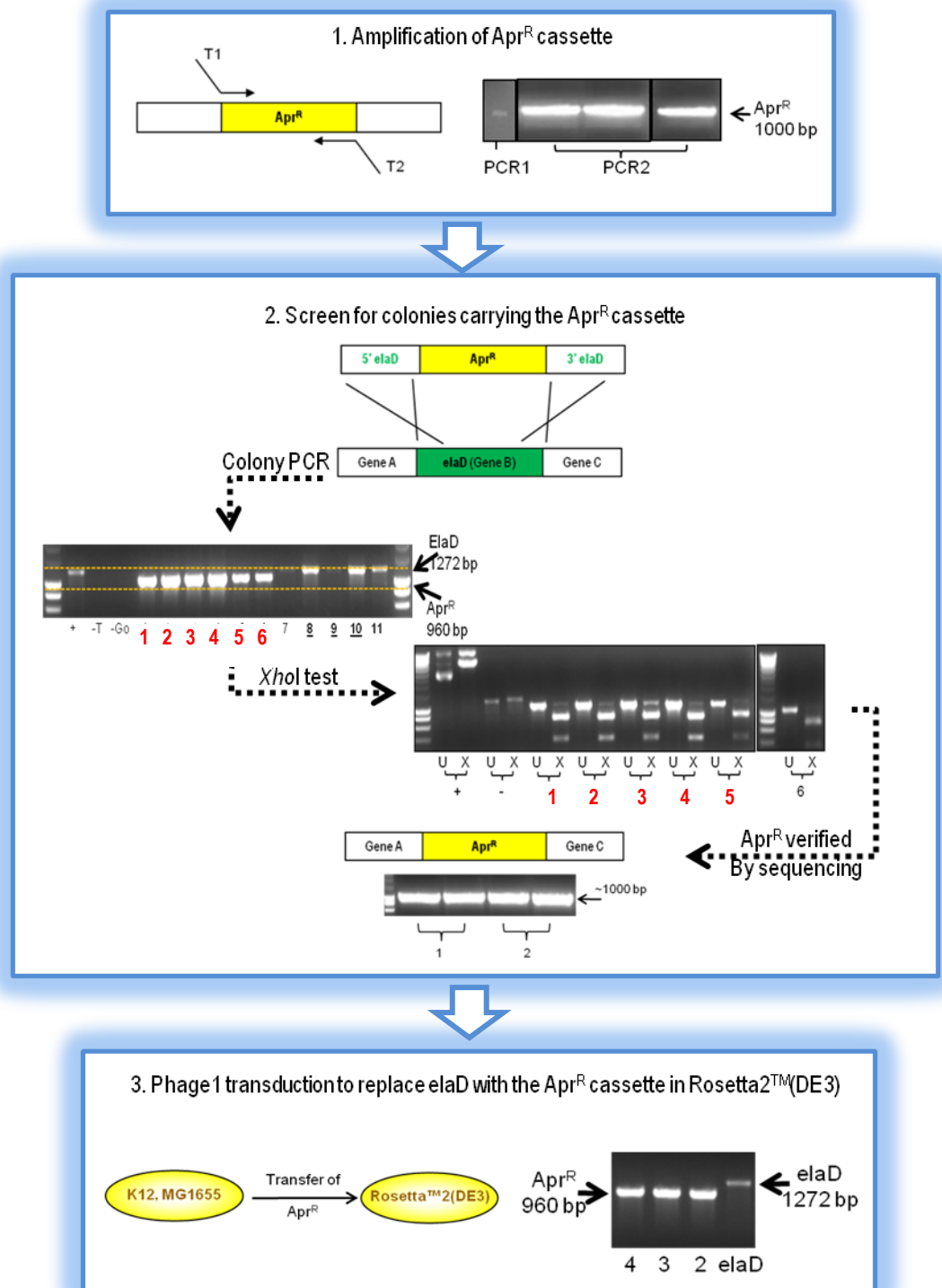
two subsequent transformations (by electroporation, 2.3.2.2). The first transformation was with pKD46 which codes for the  $\lambda$  Red proteins responsible for homologous recombination. The newly transformed JIG-182 cells were grown, induced to express  $\lambda$  Red proteins, made competent, then transformed with DNA coding for the Apr<sup>R</sup> cassette from Step 1 (details in 2.3.2). The resulting colonies were screened for JIG182 $\Delta$ elaD<sup>Apr</sup> using colony PCR (2.2.6) then positive colonies were subsequently screened using a *Xho*I test. The *Xho*I test was carried out to eliminate the possibility that an error may have occurred during the homologous recombination step resulting in the insertion of the antibiotic cassette into a random region of the genomic DNA in K12. The Apr<sup>R</sup> sequence is cut by *Xho*I and elaD is not.

The resulting DNA from both screens was separated on a 1% (w/v) agarose gel (Fig.4.7 Step 2). Clones 1 - 6 showed positive for the Apr<sup>R</sup> cassette as there was a 960 bp band. A parental Rosetta<sup>TM</sup>2(DE3) colony was included with elaD (+) amplified for comparison, - template DNA (-T) and -GoTaq (-Go). Clones 8 and 10 had a band closer to elaD (1272 bp) and there was no PCR product in clones 7 and 9.

The same DNA, for clones 1 – 6, amplified in colony PCR was then *Xho*I digested (2.2.6). *Xho*I digested (X) products and uncut Apr<sup>R</sup> DNA (U) for clones 1 – 6, as well as DNA carrying a *Xho*I site (+) and DNA

without a *XhoI* site (-) are shown in Fig.4.7 Step 2, as indicated. It is clear from the gel that clones 1 - 5 proved positive for the *XhoI* restriction sites, so amplified DNA from clones 1 and 2 were gel purified and submitted for sequencing. Both clones were JIG $\Delta$ elaD<sup>apr</sup>, verified by sequencing and clone 1 was used for the next step.

The third step involved the transfer of the *elaD* knock-out characteristic from JIG $\Delta$ elaD<sup>apr</sup> to Rosetta™2(DE3), to make Rosetta™2(DE3) $\Delta$ elaD<sup>apr</sup>. This was done using P1 transduction to carry out the homologous recombination step instead of the  $\lambda$  Red proteins (detailed in 2.2.10). A colony PCR was repeated to screen Rosetta™2(DE3) for the Apr<sup>R</sup> cassette and the resulting DNA was separated on a 1% (w/v) agarose gel and *elaD* was included for comparison (Fig.4.7 Step 3). Clones 4 - 2 showed positive for the Apr<sup>R</sup> cassette and therefore were Rosetta™2(DE3) $\Delta$ elaD<sup>apr</sup>, verified by sequencing.



**Fig.4.7**



**Fig. 4.7** Shows the results from each of the 3 main steps in developing the *elaD* knock-out in *E. coli* substrain Rosetta™2(DE3) (see text for details). Each step includes cartoons from Fig.4.6 to illustrate the procedure carried out. In Step 1 PCR was used to amplify the *Apr<sup>R</sup>* sequence. Purified *Apr<sup>R</sup>* DNA from Step 1 was then used for homologous recombination in Step 2. Resulting colonies were screened using colony PCR including *elaD* (+) to help with comparing the bands and two PCR controls, - template DNA (-T) and -GoTaq (-Go). *XhoI* digestion (X) of clones 1 – 6, as well as DNA carrying a *XhoI* site (+) and DNA without a *XhoI* site (-) identified as clones 1-4 being positive. Amplified DNA from clones 1 and 2 were gel purified and presence of the *Apr<sup>R</sup>* cassette was verified by sequencing, as indicated. The *Apr<sup>R</sup>* cassette was transferred to Rosetta™2(DE3), using P1 transduction. A colony PCR was repeated to screen Rosetta™2(DE3) for the *Apr<sup>R</sup>* cassette, including a lane with *elaD* for comparison. Clones 4, 3 and 2 showed positive for the *Apr<sup>R</sup>* cassette.

#### **4.2.3 Expression of (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP recombinant proteins in Rosetta™2(DE3)ΔelaD<sup>apr</sup> to serve as substrates for the DUB activity**

Rosetta™2(DE3)ΔelaD<sup>apr</sup> was generated to determine if elaD may be responsible for the DUB activity already described in the parental substrain of *E. coli*, in chapter 3. To investigate this possibility the two fusion proteins, (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, demonstrated previously to be cleaved by the DUB activity in the parental substrain were expressed in Rosetta™2(DE3)ΔelaD<sup>apr</sup>. This could provide convincing evidence for elaD as the DUB activity if the removal of elaD was concomitant with the absence of DUB activity. Additionally, fusion proteins which previously proved resistant to DUB cleavage were also expressed in Rosetta™2(DE3)ΔelaD<sup>apr</sup>. This was to investigate any differences in banding patterns of proteins analysed by SDS-PAGE, to see if elaD was responsible for any non-specific cleavage of fusion proteins.

Rosetta™2(DE3)ΔelaD<sup>apr</sup> and the parental substrains were transformed using the pHUE vector, sub-cloned with constructs to express the proteins: (His)<sub>6</sub>-Ub-EGFP (36 kDa), (His)<sub>6</sub>-SUMO1-LmrC (77 kDa), (His)<sub>6</sub>-NEDD8-LmrC (75 kDa) (A.I-III respectively) or the 75 kDa proteins, (His)<sub>6</sub>-Ub-LmrC (WT) and mutants, GA,GP and GGP (B).

Parental and Rosetta™2(DE3)ΔelaD<sup>apr</sup> substrains were grown as described in 2.1.1 and protein purified (2.4.1) and electrophoresed on a 5-20% (w/v) gradient gel and either Coomassie stained (Fig.4.8.B) or transferred on to nitrocellulose membrane and probed for either hexahistidine (Fig.4.8.A.I, III and C) as described in 2.4.6.2 or SUMO1 (Fig.4.8.A.II) using methods described in 2.4.6.5. Intact fusion proteins are as indicated (ζ). Both (His)<sub>6</sub>-Ub-EGFP and (His)<sub>6</sub>-Ub-LmrC fusion proteins had an 11 kDa DUB cleavage product, (His)<sub>6</sub>-Ub in the parental substrain, however, as noted previously, this DUB cleavage product was not present in the Rosetta™2(DE3)ΔelaD<sup>apr</sup>. This lack of DUB cleavage is clearly shown for (His)<sub>6</sub>-Ub-LmrC expressed in Rosetta™2(DE3)ΔelaD<sup>apr</sup>, in the Coomassie stained gel, Fig.4.8.B. The absence of DUB activity in Rosetta™2(DE3)ΔelaD<sup>apr</sup> provides the most convincing evidence so far, for elaD being responsible for the DUB activity observed in *E. coli* substrain Rosetta™2(DE3). There was no change in the banding pattern for (His)<sub>6</sub>-NEDD8-LmrC and (His)<sub>6</sub>-SUMO1-LmrC when comparing proteins expressed in Rosetta™2(DE3)ΔelaD<sup>apr</sup> with those expressed in the parental substrain. Many smaller proteins still co-purified with (His)<sub>6</sub>-SUMO1-LmrC when elaD had been knocked out which shows that elaD was not associated with this non-specific protease activity. Likewise, the 12 kDa band resulting from non-specific (NS) cleavage of GA, GP and GGP, 10 residues C-terminal to ubiquitin was present in the parental substrain as

well as in Rosetta™2(DE3)ΔelaD<sup>apf</sup>, which shows that in addition elaD was not associated with this non-specific protease activity.

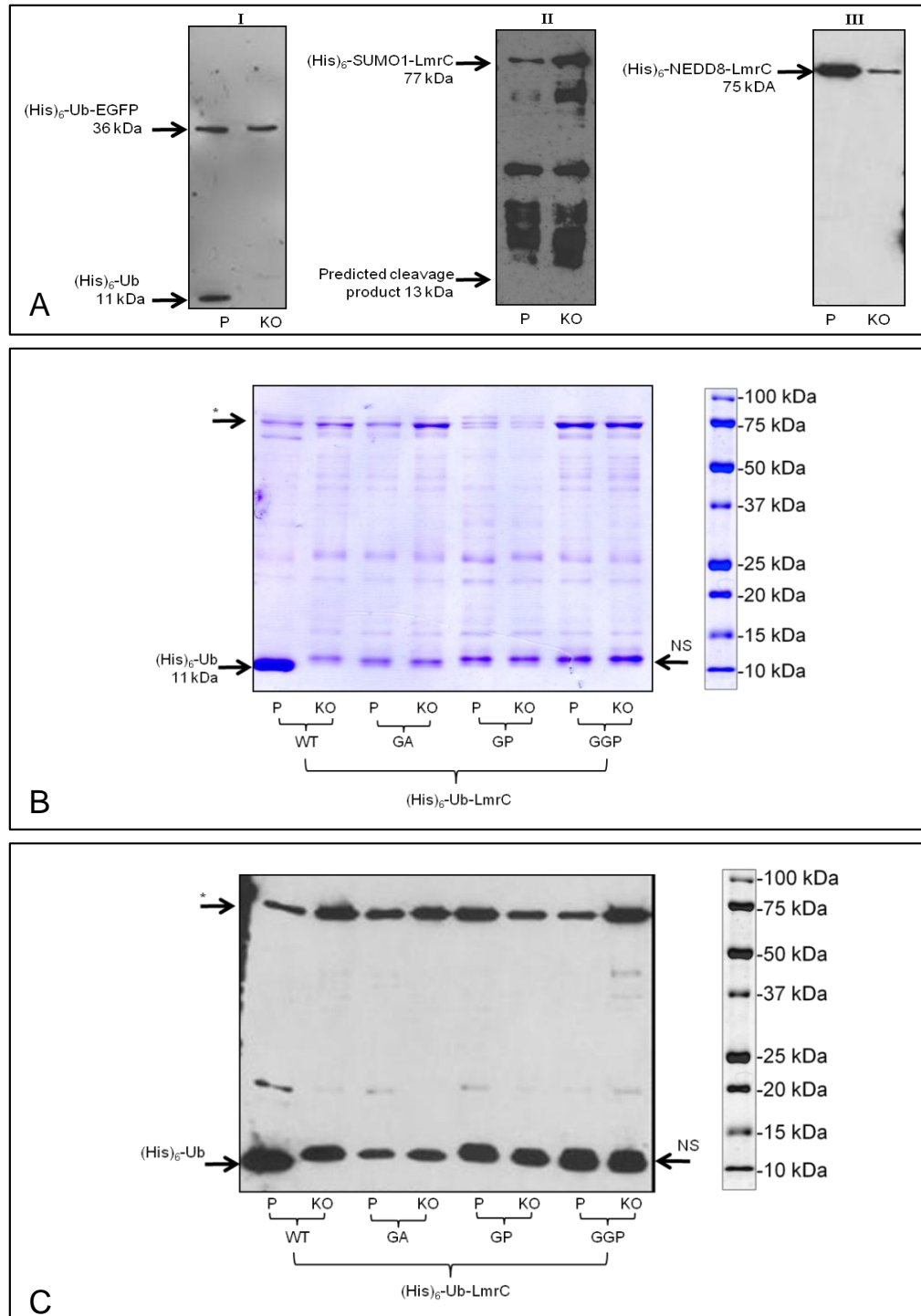


Fig.4.8

**Fig. 4.8** Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>Apr</sup> (KO) and the parental (P) cell line which expressed elaD in the genome, were transformed using the pHUE vector, sub-cloned with constructs to express the proteins: (His)<sub>6</sub>-Ub-EGFP (36 kDa), (His)<sub>6</sub>-SUMO1-LmrC (77 kDa), (His)<sub>6</sub>-NEDD8-LmrC (75 kDa) (A.I-III respectively) or the 75 kDa proteins, (His)<sub>6</sub>-Ub-LmrC (WT) and mutants, GA,GP and GGP (B and C). The resulting protein was purified using cobalt affinity chromatography, electrophoresed on a 5-20% (w/v) gradient gel and either Coomassie stained (B) or transferred on to nitrocellulose membrane and probed for either hexahistidine (A.I, III and B) or SUMO1 (A.II). Intact fusion proteins are as indicated ( $\zeta$ ). Both (His)<sub>6</sub>-Ub-EGFP and (His)<sub>6</sub>-Ub-LmrC fusion proteins had an 11 kDa DUB cleavage product, (His)<sub>6</sub>-Ub in the parental substrain, however, this DUB cleavage product was not present in the elaD knock-out substrain evident in both Coomassie stained gel (B) and the Western blot (C). This absence of DUB activity in Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>Apr</sup> is convincing evidence for elaD being a strong candidate for the DUB activity observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). There was no change in the banding pattern for (His)<sub>6</sub>-NEDD8-LmrC and (His)<sub>6</sub>-SUMO1-LmrC when comparing proteins expressed in Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>Apr</sup> with those expressed in parental *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) (as many, smaller proteins still co-purified with (His)<sub>6</sub>-SUMO1-LmrC when elaD had been knocked out this shows that elaD was not associated with this non-specific protease activity). The 12 kDa band resulting from non-specific (NS) cleavage of GA, GP and GGP, 10 residues C-terminal to ubiquitin was present in the parental substrain as well as in the elaD knock-out, which shows that elaD was not associated with this non-specific protease activity.

### **4.3 Discussion**

The main aim of the work discussed in this chapter was to identify a candidate for the DUB activity observed in *E. coli* substrain Rosetta™2(DE3), described in Chapter 3. DUB candidates, EF-Tu (TufA and TufB proteins) and *elaD* were chosen as likely candidates from the literature. These DUB candidates were then sub-cloned to eventually express GST-tagged DUB candidate proteins *E. coli* substrain XL10Gold. After optimising GST-*elaD* protein expression and purification, all candidates were then investigated to see if they could bind to ubiquitin.

The specific DUB cleavage of (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP by the DUB activity observed in Chapter 3 meant that the DUB activity could recognise and bind to ubiquitin in some way in order to cleave the ubiquitin fusion proteins. Therefore an assay was chosen to screen DUB candidates, TufA, TufB and *elaD* for the ability to bind ubiquitin. As well as the DUB candidates, eukaryotic EF-1 $\alpha$ 2 was also included in this study, to compare both eukaryotic and prokaryotic elongation factors and GST as a control. It is clear from the results of this ubiquitin-Sepharose binding assay (Fig.4.5) that *elaD* was a good candidate for the DUB activity in *E. coli* substrain Rosetta™2(DE3) compared with TufA and TufB, hence the subsequent focus on *elaD*.

It is interesting however, that neither, prokaryotic or eukaryotic elongation factors, EF-Tu ('TufA' and 'TufB' proteins) or EF1 $\alpha$ 2, respectively bound to ubiquitin-Sepharose. A. Ciechanover and colleagues reported DUB activity, *in vitro* by EF1 $\alpha$  and EF-Tu with a ubiquitinated histone, H2A as a substrate. Perhaps the DUB activity observed was actually due to *elaD* co-purifying with the elongation factors as they were purified using gel filtration. *elaD* and EF-Tu are close in mass, 45 kDa and 44 kDa respectively, and therefore it is plausible that *elaD* may have been the enzyme responsible for the DUB activity observed.

These results provide evidence for *elaD* being a strong candidate for the DUB activity observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). The next step was to ensure that *elaD* was responsible for the DUB activity observed initially with (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, using a gene knock-out technique. If *elaD* was knocked out in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) and the DUB activity was lost, then this promised to provide the strongest evidence for *elaD* as the DUB activity. Of course, this approach made the assumption that the protein targeted to be knocked out (*elaD*) was responsible for the phenotype (DUB activity). It was also possible that the target protein could instead, be a protein which influenced the expression of a protein responsible for the phenotype. For example, *elaD* could have been a transcription factor



which once activated could have lead to the eventual translation of a gene coding for the true protein responsible for the DUB activity. If this was the case, then knocking out *elaD* would also abolish DUB activity. However, in this case, as it is known that *elaD* has DUB activity, then it is safe to assume that if no DUB activity occurred when *elaD* was knocked out, then *elaD* must be the protein responsible for the DUB activity<sup>14</sup>.

In an attempt to demonstrate *elaD* was responsible for the DUB activity observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3), first, an *elaD* gene knock-out in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was created. Second, this Rosetta<sup>TM</sup>2(DE3) $\Delta$ *elaD*<sup>apr</sup> was then used to express all fusion proteins that were investigated in Chapter 3, in comparison with the parental strain. (His)<sub>6</sub>-Ub-EGFP, (His)<sub>6</sub>-SUMO1-LmrC, (His)<sub>6</sub>-NEDD8-LmrC, (His)<sub>6</sub>-Ub-LmrC and mutants were analysed by SDS-PAGE and Western blotted (Fig.4.8). When (His)<sub>6</sub>-Ub-EGFP (Fig.4.8.A) and (His)<sub>6</sub>-Ub-LmrC (Fig.4.8.B) were expressed in Rosetta<sup>TM</sup>2(DE3) $\Delta$ *elaD*<sup>apr</sup> compared with the parental substrain the band for the cleavage product, (His)<sub>6</sub>-Ub, was absent. This provides the most robust evidence, so far for *elaD* being responsible for the DUB activity that we have observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). Additionally, it was demonstrated that *elaD* was not responsible for non-specific protease activity observed with (His)<sub>6</sub>-SUMO1-LmrC and at the second, non-specific site

10aa C-terminal to the DUB site within GA, GP and GGP proteins. The banding patterns for proteins purified from cells expressing (His)<sub>6</sub>-SUMO1-LmrC, GA, GP and GGP were the same when expressed in Rosetta<sup>TM</sup>2(DE3)ΔelaD<sup>apr</sup> compared with the parental substrain (Fig.4.8.A), confirming that this resulted from non-specific cleavage as expected.

In Chapter 3, it was demonstrated that the DUB activity observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) cleaved (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, specifically in a way comparable to eukaryotic DUBs. The work discussed in the current chapter has taken a candidate approach, using ubiquitin binding and gene knock-out techniques to successfully identify elaD as the DUB activity which we have observed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). Next, having identified the DUB activity as elaD it was possible to build upon the work in Chapter 3, in further characterising elaD function.

## 5 Further characterisation of *elaD* DUB activity

### 5.1 Introduction

The work for this study began by confirming that a DUB activity observed by other members of our group, in an *E. coli* DE3 strain was reproducible in the *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). This is described in Chapter 3. Once this activity was confirmed an initial characterisation of the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was carried out. By expressing (His)<sub>6</sub>-Ub-LmrC fusion proteins, GA, GP and GGP mutated around the DUB site it was revealed that the prokaryotic DUB was very similar to both yeast and mammalian DUBs in that it did not cleave GP, i.e. post proline at position 76 of ubiquitin. However the catalytic mechanism of the prokaryotic DUB was more similar to yeast DUBs than mammalian DUBs in that it could cleave GA. It was further demonstrated, using fusion proteins expressed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) that the DUB activity showed a preference for ubiquitin as it cleaved (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, yet it did not cleave (His)<sub>6</sub>-SUMO1-LmrC or (His)<sub>6</sub>-NEDD8-LmrC, *in vivo*.

In work discussed in Chapter 4 the main focus was to identify the enzyme responsible for the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) using a candidate approach. Out of the two candidates considered for the DUB activity, the *E. coli* elongation factor,

GST-EF-Tu and GST-elaD, a ubiquitin-Sepharose pull down assay succeeded in demonstrating that only GST-elaD bound to ubiquitin. Subsequently, when elaD was knocked out of *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) this was also seen to abolish DUB activity with (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP as substrates. Therefore this work successfully identified elaD as the enzyme responsible for the DUB activity against these substrates in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3).

The preference shown by GST-elaD for ubiquitin as a substrate is interesting as *E. coli* do not express ubiquitin or a ubiquitin homologue which functions as a protein tag. *E. coli* does have ubiquitin homologues, molybdopterin converting factor, subunit 1 (MoaD) and thiamine biosynthesis protein S (ThiS), involved in Molybdenum cofactor biosynthesis and the biosynthesis of thiamine, respectively<sup>152, 153</sup>. Although MoaD and ThiS are both activated by E1 like enzymes, MoeB and ThiF (respectively), once activated instead of functioning as protein tags, MoaD and ThiS provide a sulphur obtained during activation for incorporation into molybdopterin and thiazole, respectively<sup>154, 155</sup>. It is possible that elaD may function to subvert the host cell system by targeting the eukaryotic ubiquitin system. Alternatively elaD may form part of a eukaryotic ubiquitin-like system in *E. coli*, yet to be discovered. In this Chapter the aim was to investigate further the specificity of GST-elaD for different types of ubiquitin

substrates. The work using deubiquitinating assays, described so far in this study has focussed on DUB activity of GST-elaD, *in vivo* with fusion proteins. The work in the current Chapter uses a variety of ubiquitin substrates to build a more detailed picture of the catalytic mechanism and specificity of GST-elaD.

In eukaryotic cells a large variety of ubiquitin ligated proteins form substrates for eukaryotic DUBs within the eukaryotic ubiquitin system. There are seven possible lysine (K) residues in ubiquitin available for another ubiquitin molecule to become covalently attached: K6, K11, K27, K29, K33 K48 and K63 (Fig.5.1.A) and are used *in vivo*. Ubiquitin links to its interaction partners which can either be a target protein (Fig.5.1.B) or another ubiquitin molecule, leading to the formation of poly-ubiquitin chains (Fig.5.1.C).

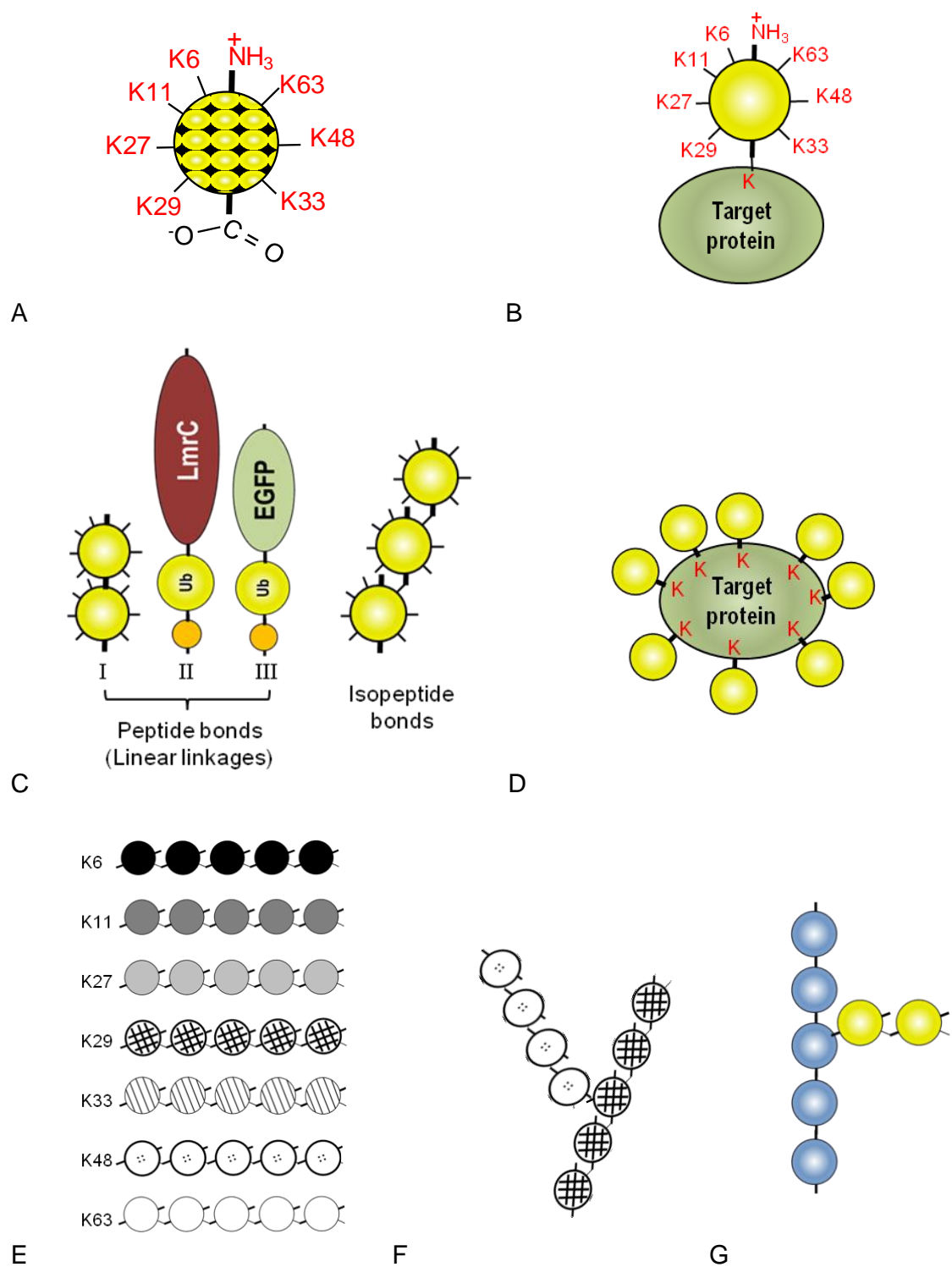
There are two main types of bond between ubiquitin proteins or between ubiquitin or another protein; the first is a linear linkage (peptide bond) and the second is an isopeptide bond. The linear linkage exists between the  $\alpha$ -amino group of the first amino acid residue of ubiquitin, methionine (M1), with the  $\alpha$ -carboxyl group (G76) on another ubiquitin (as occurs in ubiquitin precursors) or another protein (as in artificial fusion proteins) (Fig.5.1.C.II-III). The linear fusion proteins, (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, used as substrates for GST-elaD in

previous chapters are shown, with orange spheres representing (His)<sub>6</sub>. The linear linkage can occur between ubiquitin precursors during translation or as was recently discovered, there is an E3 ligase complex, the linear ubiquitin chain assembly complex (LUBAC) which assembles linear ubiquitin chains on target proteins<sup>78</sup> (Fig.5.1.CI).

Isopeptide bonds are found between ubiquitin and target proteins, ubiquitin dimers or ubiquitin polymers. Isopeptide bonds occur post translationally using the  $\epsilon$ -amino group of lysine on one ubiquitin and the  $\alpha$ -carboxyl group of G76 of ubiquitin. An isopeptide bond can occur using many lysine residues within the target protein to produce a multiply ubiquitinated protein (Fig.5.1.D) or poly-ubiquitin chains. Ubiquitin chains can form a single isomer (Fig.5.1.E), or a mixed chain with different lysine residues used which can form a fork. A forked chain is shown in (Fig.5.1.F) in which a K29-linked ubiquitin chain forms a bond with two other ubiquitins using K29 and K48, causing a fork. Additionally it is possible for heterologous chains to form, for example, SUMO2 chains (blue spheres) have been shown to become ubiquitinated with a ubiquitin chain (Fig.5.1.G).

The work discussed in this Chapter, concentrates on building a clearer picture of the nature of the catalytic activity of GST-elaD using a wider range of ubiquitin substrates. Linear linkages between two ubiquitin

molecules, isopeptide bonds between ubiquitin of two to four units long, in mixed length poly-ubiquitin chains, in single poly-ubiquitin chains exclusively using one of all seven possible lysine linkages were used. Finally possible binding mutants were made and used in a ubiquitin-Sepharose pull-down assay to ascertain the regions within GST-elaD responsible for binding to ubiquitin (seen previously 4.2.4).



**Fig.5.1**



**Fig. 5.1** Ubiquitin is shown as a yellow sphere with the carboxyl group of the C-terminus and the  $\alpha$ -amino group of the N-terminus. The ubiquitin  $\alpha$ -amino group and 7 lysine residues (K) which have an  $\epsilon$ -amino group, are available to form linkages with other ubiquitin proteins (A). Ubiquitin forms a bond with a second protein which can either be a target protein (B) or another ubiquitin molecule which can also form ubiquitin chains shown in (C). This bond can occur using the  $\alpha$ -carbonyl carbon on ubiquitin and nitrogen from either the  $\alpha$ -amino group or the  $\epsilon$ -amino group of the second protein. If ubiquitin is connected to the  $\alpha$ -amino group of another protein then this is a peptide bond, known as linear linkage and occurs both during translation (C.II-III), or post translationally. The linear fusion proteins, (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP, used as substrates for *elaD* in previous chapters are shown, with orange spheres representing (His)<sub>6</sub>. The ubiquitin E3 ligase, linear ubiquitin chain assembly complex (LUBAC) builds chains post translationally (C.I). Alternatively, if ubiquitin is connected to the  $\epsilon$ -amino group of another protein then this is an isopeptide bond which is only formed post translationally. An isopeptide bond can occur using many lysines within the target protein to produce a multiply ubiquitinated protein (D) or ubiquitin chains. Ubiquitin chains can form a single isomer (E), or a mixed chain with different lysine residues used which can form a fork. Just one example of a forked chain is shown in (F) in which a K29 linked ubiquitin chain forms a bond with two other ubiquitins using K29 and K48, causing a fork. Additionally it is possible for heterologous chains to form, for example, SUMO2 chains (blue spheres) have been shown to become ubiquitinated with a ubiquitin chain (G).

## 5.2 Results

### 5.2.1 An *in vitro* DUB assay using (His)<sub>6</sub>-Ub-LmrC as a substrate for GST-elaD

The fusion protein (His)<sub>6</sub>-Ub-LmrC (Fig.5.2) was used as a substrate for GST-elaD. (His)<sub>6</sub>-Ub-LmrC was demonstrated to be cleaved, *in vivo* by elaD in previous chapters was not completely processed by elaD (see

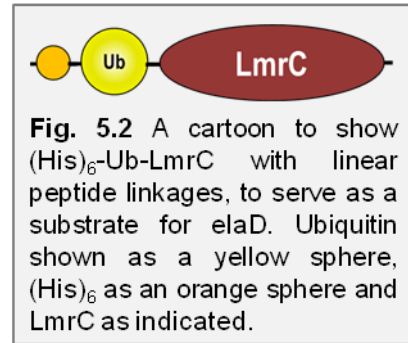


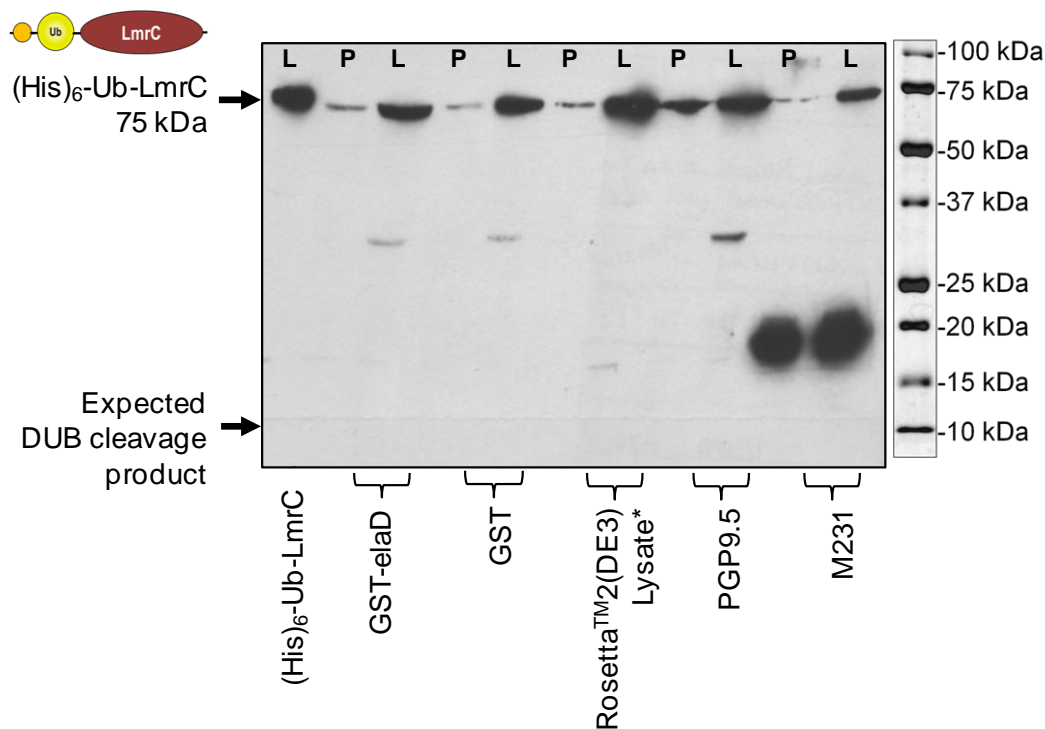
Fig.3.2). This led to the speculation that as LmrC is thought to form homo-dimers, perhaps cleaved LmrC and intact fusion proteins formed dimers which prevented further DUB activity<sup>156</sup>.

To test this idea intact fusion protein was used as a substrate for GST-elaD. The intact (His)<sub>6</sub>-Ub-LmrC substrate was induced in *L. Lactis*, a bacterial strain which does not express elaD and only produces the intact form of the fusion protein. The intact (His)<sub>6</sub>-Ub-LmrC used in the assay was either purified (P) protein or used as a lysate containing unpurified recombinant (His)<sub>6</sub>-Ub-LmrC protein (L). Two DUBs were included as positive controls for DUB activity, a eukaryotic DUB, protein gene product 9.5 (PGP9.5) and a viral DUB, M231. Cell lysate containing unpurified recombinant proteins: GST-elaD, GST, GST-PGP9.5 and purified M231 protein were used in the DUB assay. A cell

lysate from Rosetta<sup>TM</sup>2(DE3) with no pHUE vector was also included, to control for any background DUB activity by intrinsically expressed elaD. This would allow a comparison to be made between the DUB activity of unpurified GST-elaD (which was in a cell lysate) with any DUB activity of intrinsically expressed elaD in the same volume of cell lysate. The amounts of enzyme or substrate used in each DUB assay reaction were produced by a 10 ml culture. The reactants were incubated (37°C) in TBS with 0.5 mM DTT (pH 7.5) for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for the hexahistidine-tag (Fig.5.3).

There was no evidence of DUB activity by any of the DUBs using the (His)<sub>6</sub>-Ub-LmrC as a substrate, *in vitro*, evident from the absence of cleavage product, (His)<sub>6</sub>-Ub (Fig.5.3). The lack of DUB activity by PGP9.5 and M231 as well as GST-elaD would seem to indicate that there may be an alternative block to DUB activity against this substrate post translationally; perhaps when (His)<sub>6</sub>-Ub-LmrC is expressed in *L. Lactis*, strong dimerisation occurs between LmrC intact fusion proteins precluding subsequent proteolysis by DUBs. This could also be the reason for the lack of background DUB activity by constitutively expressed elaD (Rosetta<sup>TM</sup>2(DE3) with no pHUE vector) (Fig.5.3). Alternatively, unlike all DUB assays carried out previously, this DUB

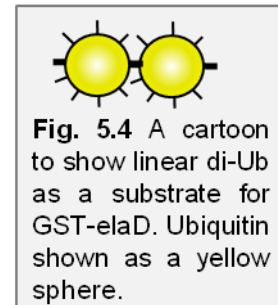
assay was carried out *in vitro*, rather than an *in vivo*. Therefore as the protein expression (of *elaD* and substrate (His)<sub>6</sub>-Ub-LmrC) and DUB activity did not occur within *E. coli*, the lack of any DUB activity (including background DUB activity) by *elaD* may have been due to the lack of an enzyme co-factor or a key protein binding partner. Additionally, the absence of background DUB activity (by intrinsically expressed *elaD*) may have been due to the level of *elaD* being so low that DUB activity was undetectable. The reason for including Rosetta<sup>TM</sup>2(DE3) with no pHUE vector was to compare the DUB activity of unpurified GST-*elaD* (in lysate) with any background DUB activity, by intrinsically expressed *elaD* in the same amount of cell lysate taken from a 10 ml culture.



**Fig. 5.3** The fusion protein, (His)<sub>6</sub>-Ub-LmrC containing linear peptide linkages between ubiquitin and LmrC was used as a substrate for GST-elaD. The DUB assay using (His)<sub>6</sub>-Ub-LmrC, expressed by *L. lactis* as a substrate for elaD was set up as follows. *E. coli* substrain Rosetta™2(DE3) cell lysate, containing unpurified proteins: GST-elaD, GST, PGP9.5 or purified proteins: USP2-core M231 were incubated (37°C) with either purified (P) or unpurified (L) (His)<sub>6</sub>-Ub-LmrC protein in TBS with 0.5 mM DTT (pH 7.5) for approximately 15 hr. PGP9.5 and M231 were included as positive controls for DUB activity. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide, gradient gel transferred on to nitrocellulose membrane and probed for the hexahistidine-tag. There was no evidence of DUB activity by any of the DUBs using the (His)<sub>6</sub>-Ub-LmrC as a substrate, *in vitro*. (His)<sub>6</sub>-Ub-LmrC was expressed by *L. lactis* and therefore intact and available for DUB cleavage. However there was no DUB activity with (His)<sub>6</sub>-Ub-LmrC by PGP9.5 and M231 as well as elaD which indicates a block to DUB activity. There is cross reactivity of the hexahistidine-tag antibody with M231 enzyme.

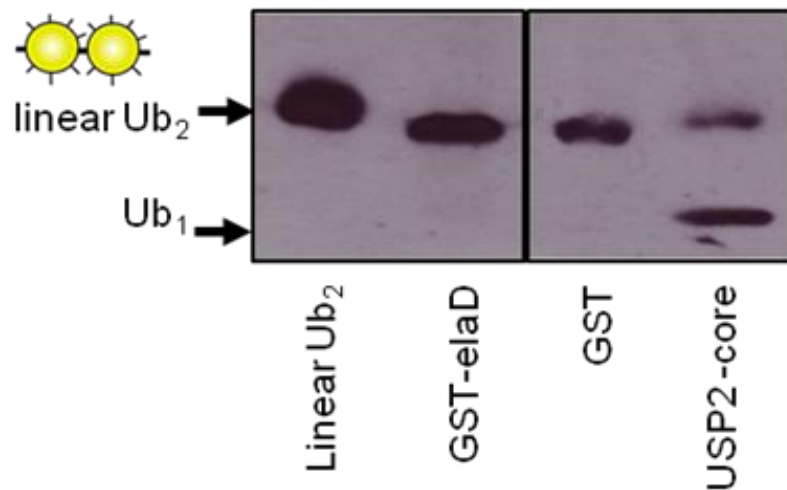
### 5.2.2 An *in vitro* DUB assay using unanchored linear di-ubiquitin as a substrate for GST-elaD

Linear di-ubiquitin was used as a substrate for GST-elaD, as an alternative linear peptide-linked protein to use *in vitro*, without complication of possible substrate dimerisation. GST-elaD (~15 µg) and GST (~15 µg) were purified using



glutathione-Sepharose 4B beads and while still attached to beads were then incubated (37°C) with 0.9 µg linear di-ubiquitin (Ub<sub>2</sub>) in TBS with 0.5 mM DTT (pH 7.5) for approximately 15 hr. USP2-core (0.4 µg) was included as a DUB to serve as a positive control. The reaction was again stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.5.5).

Interestingly, linear di-ubiquitin was not deconjugated by immobilised GST-elaD, but was deconjugated by USP2-core, *in vitro*. The absence of DUB activity by GST-elaD with Ub<sub>2</sub> was surprising as the bonds between ubiquitin and LmrC or another ubiquitin, in this case were both linear, i.e. peptide linkages.

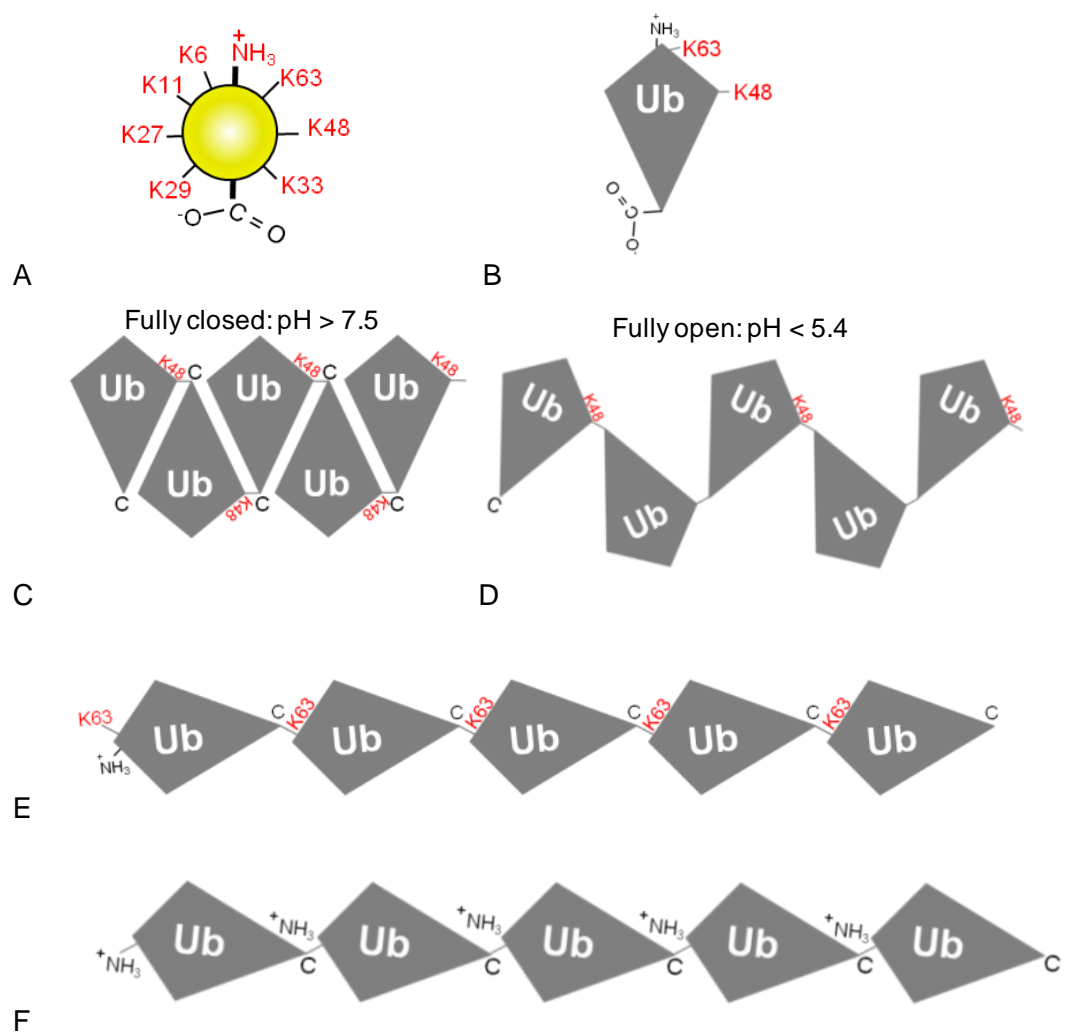


**Fig.5.5** Linear di-ubiquitin was used as a substrate for GST-elaD. Immobilised GST-elaD (15  $\mu$ g) and GST were incubated (37°C) with 0.9  $\mu$ g linear di-ubiquitin (Ub<sub>2</sub>) in TBS with 0.5 mM DTT (pH 7.5) for approximately 15 hr. USP2-core (0.4  $\mu$ g) was included as a DUB to serve as a positive control for DUB activity. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (B). Linear di-ubiquitin (Ub<sub>2</sub>) was not cleaved by elaD but was cleaved by USP2-core.

The next substrates considered for GST-elaD were anchored K63-linked and K48-linked tetra-ubiquitin. However, first, this would be good point to consider the different conformations of the different isomers of poly-ubiquitin chains. Fig.5.6 shows cartoons of ubiquitin as either the spherical shape introduced earlier (Fig.5.6.A), or as a kite shape which allows the conformation of ubiquitin chains to be shown. At the time of writing, there was only information in the literature about the conformation structure of K48, K63 and linear ubiquitin dimers, which can be extrapolated to poly-ubiquitin chains. It has been shown that K48-linked poly-ubiquitin chains adopt a closed conformation with increasing pH, being fully closed when pH is greater than 7.5, and is fully open when pH is lower than 5.4 (Fig.5.6.D)<sup>16</sup>. The K63-linked ubiquitin chain has been demonstrated to form a more open conformation than the relaxed form of K48-linked ubiquitin chain, with a strong similarity to linear ubiquitin (Fig.5.6.F)<sup>19</sup>.

Knowing that GST-elaD cleaves a linear fusion protein which contains ubiquitin (His<sub>6</sub>-Ub-LmrC) would lead one to speculate that GST-elaD might show a preference for ubiquitin polymers with an open conformation, such as a linear linkage or possibly a K63-linked ubiquitin chains. As there was no evidence for deconjugation of linear di-ubiquitin by GST-elaD so next, the K48 and K63 linked tetra-ubiquitin were evaluated as substrates for GST-elaD.





**Fig. 5.6** A cartoon to show the different conformations of some of the ubiquitin chain isomers. Ubiquitin can be shown simplistically as a yellow sphere (A) as well as a kite shape which is closer to the ubiquitin structure and is useful for showing conformations of ubiquitin within different chain isomers. At the time of writing, there is only information in the literature about the conformation of K48, K63 and linear ubiquitin dimers, which can be extrapolated to poly-ubiquitin chains. K48-linked ubiquitin chains are able to adopt a closed conformation and with increasing pH the conformation opens (D)<sup>16</sup>. The K63-linked ubiquitin chain forms a more open conformation than the open form of K48-linked ubiquitin chain, with a strong similarity to linear ubiquitin (F)<sup>19</sup>.

### **5.2.3 K63 and K48 isopeptide bonds within unanchored tetra-ubiquitin as substrates for GST-elaD**

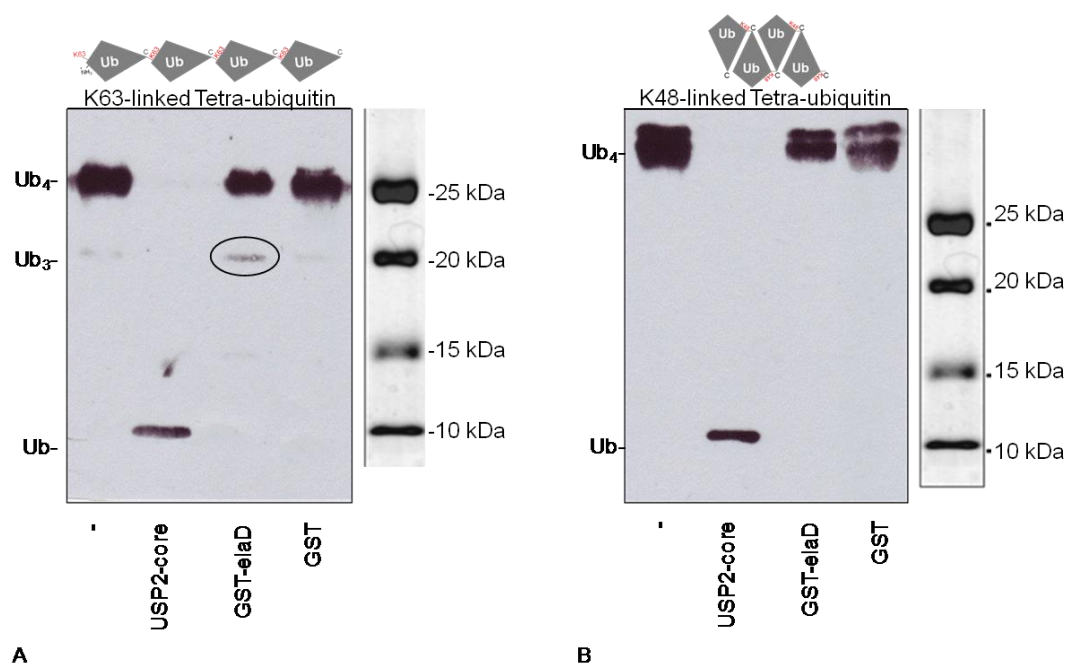
The K63 and K48 isomers of tetra-ubiquitin were investigated next as substrates for GST-elaD *in vitro*. GST-elaD and GST were purified using Glutathione-Sepharose 4B beads and after washing in TBS, whilst still immobilised on the beads GST-elaD (~15 µg) and GST (~15 µg) were incubated with 0.25 µg tetra-ubiquitin in TBS (1 mM DTT), 37°C for 15 hr. The reaction was stopped using gel loading buffer and the samples were electrophoresed on a 5-20% (w/v) acrylamide gel, transferred to nitrocellulose membrane and probed for ubiquitin (Fig.5.7). There is evidence that GST-elaD deconjugated tetra-ubiquitin to tri-ubiquitin but there were no bands representing di-ubiquitin or mono-ubiquitin on the blot so it was not complete deconjugation as with the control USP2-core. It should be noted that the K63-linked tetra-ubiquitin was not pure as traces of tri-ubiquitin were present in the control sample (lane 1 of Fig.5.7.A). However, more tri-ubiquitin is present in the reaction containing GST-elaD (lane 3 of Fig.5.7.A) than in the control sample.

There was no deconjugation by GST-elaD of K48-linked tetra-ubiquitin which would indicate that GST-elaD shows a preference for a K63-linked tetra-ubiquitin under comparable conditions. The lack of deconjugation of K48-linked tetra-ubiquitin by GST-elaD may be due to

the closed conformation assumed by this poly-ubiquitin chain isomer at the pH at which the assay was carried out (pH 7.5)<sup>16</sup>. As GST-elaD cleaves ubiquitin containing substrates with linear linkages presumably indicating a requirement by GST-elaD for the ubiquitin containing substrate to adopt an 'open' conformation to reveal necessary binding epitopes.

The discovery that GST-elaD deconjugated K63-linked tetra-ubiquitin but not K48-linked tetra-ubiquitin was an interesting finding as it appeared that GST-elaD preferred the open conformation of K63. elaD is a bacterial enzyme which is now known from previous work in this study to prefer ubiquitin as a substrate over NEDD8 or SUMO1, yet *E. coli* does not express a homologue of ubiquitin. Now this evidence shows that GST-elaD not only deconjugates isopeptide bonds between ubiquitin but apparently prefers the K63-linked tetra-ubiquitin isomer over K48. K63-linked isomers of ubiquitin chains are associated with signalling pathways in eukaryotic cell systems. This leads to the speculation that GST-elaD may have a function to overthrow, or manipulate a host cell signalling system, assuming it is a secreted bacterial effector see discussion.

Having demonstrated catalytic activity by GST-elaD *in vitro*, the next step in this investigation was to make a GST-elaD catalytic mutant in order to confirm specificity of activity.



**Fig. 5.7** GST-elaD (15  $\mu$ g) and GST, whilst still attached to glutathione-Sepharose 4B beads and USP2-core (0.4  $\mu$ g) used as a DUB to serve as a positive control for DUB activity, were incubated (37°C) with 0.25  $\mu$ g tetra-ubiquitin (Ub<sub>4</sub>), K63 (A) and K48 (B) for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. Only K63-linked Ub<sub>4</sub> was deconjugated to Ub<sub>3</sub> by elaD as indicated.

## 5.2.4 Construction of a GST-elaD catalytic mutant, GST-C313S for comparison with GST-elaD in *in vitro* DUB assays

### 5.2.4.1 Designing the GST-elaD catalytic mutant, GST-C313S

Previously (4.2.1), when planning the GST-elaD construct, the sequence for genomic elaD from *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was sequenced, then aligned with the elaD sequence in *E. coli* substrain K12. Both elaD sequences were found to be identical for both substrains of *E. coli*. A GST-elaD mutant was designed to remove the catalytic ability of GST-elaD by targeting the cysteine of the catalytic triad. Fig.5.8 shows the catalytic triad in SENP8 (*Human* GI: 33942066) compared with the elaD homologue, from two lab strains of commensal bacteria (*E. coli* Rosetta<sup>TM</sup>2(DE3) sequenced in this study) and *E. coli* K12 (GI: 16130204), as well as elaD from pathogenic bacteria, (*E. coli*\_O157 GI: 15832411). The catalytic triad and oxyanion stabilising groups for SENP8, elaD in O157 and elaD in K12 were highlighted by A. Catic<sup>14</sup>. The elaD sequence in Rosetta<sup>TM</sup>2(DE3), established earlier in this study (4.2.1) as identical to elaD in K12, thus has the same highlighted regions in Fig.5.8.

The catalytic triad for these cysteine proteases, shown in red, is formed by: cysteine (C), histidine (H) and asparagine (N) or aspartate (D), in elaD asparagine is used. Glutamine (Q), shown in aqua, provides the

oxyanion stabilising group, of which there are two in elaD. The cysteine targeted for mutation in elaD from Rosetta<sup>TM</sup>2(DE3), the 313<sup>th</sup> residue, was converted to serine (S) and therefore the catalytic mutant for GST-elaD is known as GST-C313S onwards. This strategy was also used by A. Catic to create a catalytic mutant of GST-elaD from *E. coli* substrain K12<sup>14</sup>. The GST-C313S mutant was generated by carrying out PCR based SDM using Pfu DNA polymerase as described in 2.6.4.

```

SENp8 -----
Rosetta2 (DE3) MMVTVVSNYCQLSQTQLSQTFAEKFTVTEELLQSLKKTALSGDEESIELLHNIALGYDKF 60
K12 MMVTVVSNYCQLSQTQLSQTFAEKFTVTEELLQSLKKTALSGDEESIELLHNIALGYDKF 60
O157 MMVTVVSNYCQLSQTQLSQTFAEKFTVTEELLQSLKKTALSGDEESIELLHNIALGYDEF 60

SENp8 -----MDPVVL 6
Rosetta2 (DE3) GKEAEDILYHIVRTPPTNETLSIIRLIKNACLKLYNLAHIATNSPLKSHDSDDLFFKKLFS 120
K12 GKEAEDILYHIVRTPPTNETLSIIRLIKNACLKLYNLAHIATNSPLKSHDSDDLFFKKLFS 120
O157 GKKAEDILYHIVRNPTNDTLSIIKLIKNACLKLYNLAHTATKHPLKSHDSNDLFFKKLFS 120

SENp8 SYMDSLRLRQSDVSLDPPSWLNDHIIGFAFEYFAN-SQFHDSCDHVSFISPEVTQFIKCT 65
Rosetta2 (DE3) PSKLMTIIGDEIPLISEKQSLSKVLLNDENNELSDGTNFWDKNRQLTTDEIACYLQKIAA 180
K12 PSKLMTIIGDEIPLISEKQSLSKVLLNDENNELSDGTNFWDKNRQLTTDEIACYLQKIAA 180
O157 PSKLMATIGDEIPLISEKQSLSKVLLNDKNNELSDGTNFWDKNRQLTTDEIACYLKIAA 180

SENp8 SNPAEIAMFLEPLDLPNKRNVFLAINDNSNQAAG-----GTEHWSLLVYLQD 111
Rosetta2 (DE3) NAKNTQVNYPTGLYVPYSTRTHLEDALNENIKSDPSWPNEVQLFPINTGGEHILVSLQKI 240
K12 NAKNTQVNYPTGLYVPYSTRTHLEDALNENIKSDPSWPNEVQLFPINTGGEHILVSLQKI 240
O157 NAKNTQVNYPTDFYLPNSNSTYLEVALNDNIIKSDPSWPKEVQLFPINTGGEHILVSLQKI 240

SENp8 KNSFFHYDSDHSRNSVHAKQVAEKLEAFLGRKGDKLAF-----VEEKAPAQNSYDCGM 166
Rosetta2 (DE3) VNKKNNKLQIKCVIFNSLRALGYDKENSLKRVINSFNS---ELMGEMSNN-NIKVHLN 296
K12 VNKKNNKLQIKCVIFNSLRALGYDKENSLKRVINSFNS---ELMGEMSNN-NIKVHLN 296
O157 VNEKNNTQQIKCIIFNSLRALGHEKENSLKRIINSFNSFNCPTRETPTNNKNITDHLTEP 300

SENp8 VICNTEALCQNFRRQQTESLLQLLTPAYITKKRGEWKDLIATLAKK----- 212
Rosetta2 (DE3) EIIFLHADLQQYLSQSCGAFVCMQAQEVIEQRESNSDSAPYTLKKNHADRFKKYSAAEQY 356
K12 EIIFLHADLQQYLSQSCGAFVCMQAQEVIEQRESNSDSAPYTLKKNHADRFKKYSAAEQY 356
O157 EIIFLHADLQQYLSQSCGAFVCMQAQEVIEQRESNSDSAPYTLKKNYADRFKKYSAAEQY 360

SENp8 -----
Rosetta2 (DE3) EIDFQHRLANRNCYLDKYGDANINHYRNLEIKHSQPKNRASGRVS 403
K12 EIDFQHRLANRNCYLDKYGDANINHYRNLEIKHSQPKNRASGRVS 403
O157 EIDFQHRLANRNCYLDKYGDANINHYRNLEIKNSHPKNRASGRVS 407

```

**Fig. 5.8** An alignment using full length protein sequences between SENP8 (*Human* GI: 33942066) and: elaD in two lab strains of commensal bacteria (*E. coli* Rosetta<sup>TM</sup>2(DE3 sequenced in this study) and *E. coli* K12 (GI: 16130204), including an elaD homologue found in bacteria which are pathogenic to humans (*E. coli* O157 GI: 15832411). The alignment was carried out using EMBL-EBI Clustal W online alignment tool. The predicted catalytic triad: histidine (H), aspartate (D) or asparagine (N) and cysteine (C) are all shown in red. The glutamine residues (Q) shown in aqua form part of the oxyanion-stabilising group with cysteine. A mutation was introduced in elaD (Rosetta<sup>TM</sup>2(DE3)) to remove the ability to carry out catalytic activity. A good target for mutation in cysteine proteases is the cysteine of the catalytic triad. Therefore the elaD catalytic mutant was made by converting cysteine, the 313th residue in elaD to serine (S) generating mutant, C313S. The catalytic mutant could then be used to compare with wild type elaD within DUB assays to be able to demonstrate that any activity observed *in vitro* would be as a result of DUB activity by elaD.

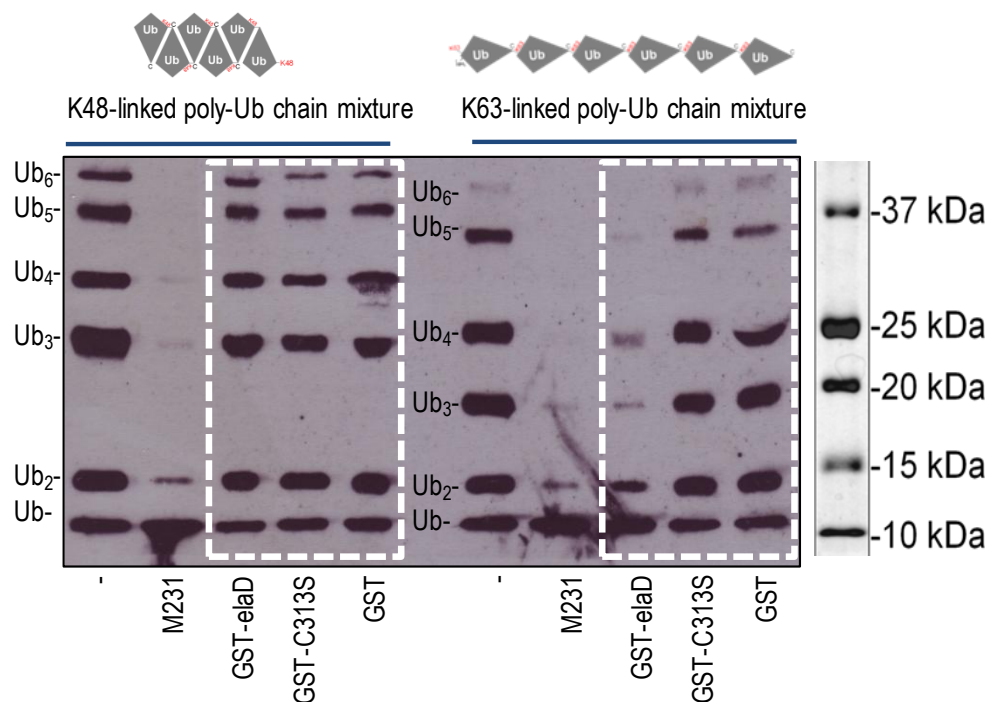


### **5.2.5 K63 and K48 linked unanchored poly-ubiquitin mixed length chains as substrates for immobilised GST-elaD including GST-C313S**

In section 5.2.2 GST-elaD showed a preference for K63-linked tetra-ubiquitin chains over K48-linked tetra-ubiquitin so the next step was to test GST-elaD for DUB activity with longer chains of ubiquitin of the same linkages. GST (~15 µg), GST-elaD (~15 µg) and the GST-tagged, newly generated catalytic mutant, GST-C313S (~15 µg) were captured on glutathione-Sepharose 4B beads, washed with 0.1% (v/v) Triton X-100 in TBS and incubated (37°C) in parallel with M231, a positive control DUB with 0.45 µg poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.5.9).

The same preference by GST-elaD for K63-linked poly-ubiquitin chains over K48-linked poly-ubiquitin chains was shown compared with tetra-ubiquitin. However, in this case a stronger DUB activity was observed. This suggests that elaD does prefer open conformations of ubiquitin such as K63-linked poly-ubiquitin chains. The GST-elaD catalytic mutant, however did not deconjugate the poly-ubiquitin chains confirming that activity by GST-elaD is a specific DUB activity. K63-linked poly-ubiquitin chains were deconjugated by GST-elaD, evident

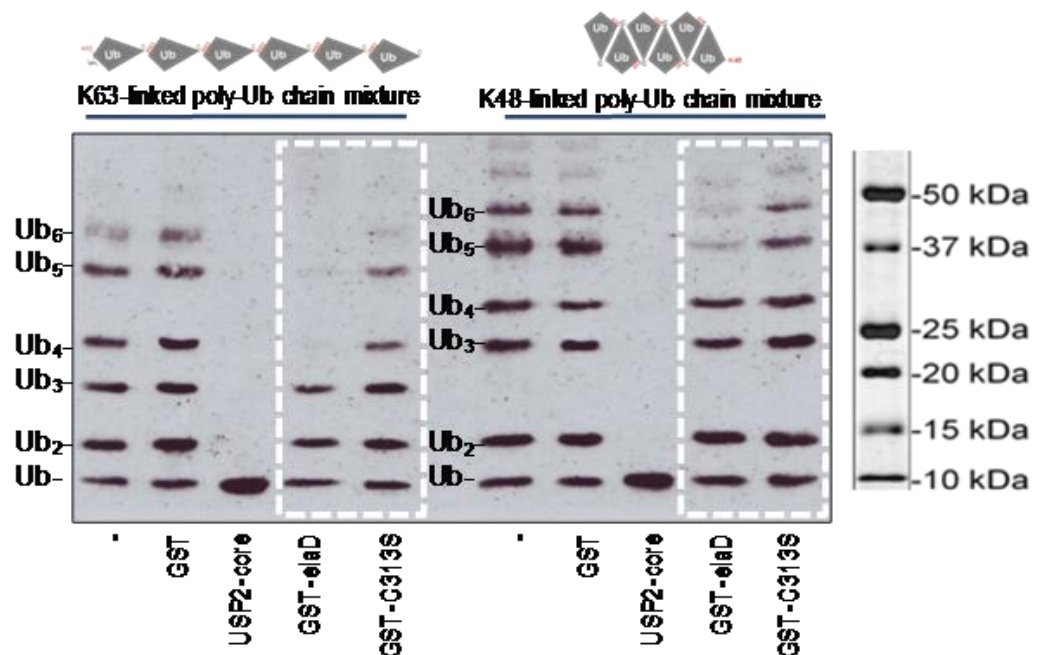
from the absence of Ub<sub>6</sub> and a barely visible band for Ub<sub>5</sub>. This was not complete deconjugation as weak bands for Ub<sub>4</sub> and Ub<sub>3</sub> were noticeable. Indeed strong bands for Ub<sub>2</sub> could indicate that GST-elaD prefers longer poly-ubiquitin chains.



**Fig. 5.9** Poly-ubiquitin chain mixtures containing either K63-linked ubiquitin or K48-linked ubiquitin were used as a substrate for GST-elaD. This was to determine if the same preference for the K63-linked form was seen with longer chains than tetra-ubiquitin. GST, GST-elaD (15  $\mu$ g) and the GST-tagged, catalytic mutant (C313S) (15  $\mu$ g) were captured on glutathione-Sepharose 4B beads, washed with 0.1% (v/v) Triton X-100 in TBS and incubated in parallel with M231, a positive control DUB with 0.45  $\mu$ g poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. C313S showed not DUB activity with either isomer of poly-ubiquitin chain mixtures. The same preference by elaD for K63-linked poly-ubiquitin chains over K48-linked poly-ubiquitin chains as was shown with tetra-ubiquitin was observed. K63-linked poly-ubiquitin chains were deconjugated, evident from the absence of Ub<sub>6</sub> and a barely visible band for Ub<sub>5</sub>. This was not complete deconjugation as weak bands for Ub<sub>4</sub> and Ub<sub>3</sub> were noticeable. Strong bands for Ub<sub>2</sub> would indicate that elaD has a preference for longer poly-ubiquitin chains.

### **5.2.6 K63 and K48 linked unanchored poly-ubiquitin mixed length chains as substrates for GST-elaD in solution**

The DUB assays using linear di-ubiquitin and K48 or K63 linked poly-ubiquitin as substrates for GST-elaD were carried out previously with GST-elaD attached to glutathione-Sepharose 4B beads. To check that the attachment of GST-elaD to beads did not alter DUB activity in any way, GST-tagged proteins were next eluted from glutathione-Sepharose 4B beads using 20 mM reduced glutathione then dialysed in 0.1% (v/v) Triton X-100 in TBS. GST (~15 µg), ~15 µg GST-elaD and ~15 µg GST-tagged, catalytic mutant (GST-C313S) were incubated (37°C) with 0.45 µg poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.5.10). It is evident from the Western blot that the same preference is shown by GST-elaD for the K63-linked poly-ubiquitin chain mixture when free from the glutathione-Sepharose 4B beads and therefore in solution, compared with when it was immobilised. Again, the longer poly-ubiquitin chains seem to be preferred by GST-elaD. Additionally the free GST-elaD appeared to show a limited deconjugation of K48-linked chains. GST-C313S and GST were not associated with any deconjugation of either chain isomers. USP2-core was included in the DUB assay as a positive control DUB fully deconjugated both poly-ubiquitin chain forms.



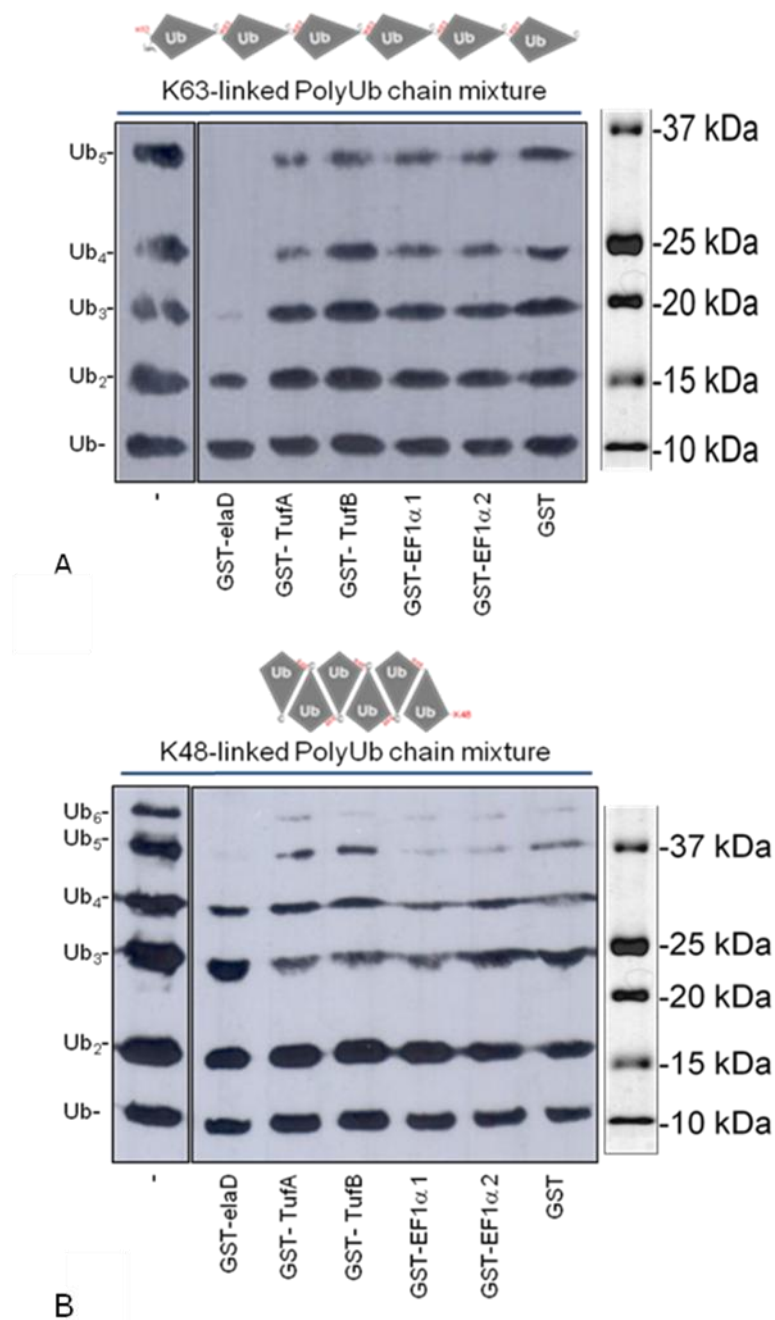
**Fig. 5.10** Poly-ubiquitin chain mixtures containing either the K63-linked ubiquitin or K48-linked ubiquitin were used as a substrate for GST-elaD, this time eluted from glutathione-Sepharose 4B beads. This assay was used to check that elaD activity was not altered by being attached to the glutathione-Sepharose 4B beads. GST, GST-elaD (15  $\mu$ g) and the GST-tagged, 15  $\mu$ g catalytic mutant (C313S) were incubated (37°C) in parallel with 0.4  $\mu$ g USP2-core, a positive control DUB with 0.45  $\mu$ g poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. The C313S mutant showed no DUB activity with either isomer of poly-ubiquitin chain mixtures. The same preference by eluted elaD was shown for K63-linked poly-ubiquitin chains over K48-linked poly-ubiquitin chains compared with when elaD was attached to glutathione-Sepharose 4B beads. K63-linked poly-ubiquitin chains were only deconjugated to Ub<sub>3</sub> this time compared with being deconjugated to Ub<sub>2</sub> in Fig.5.9. The eluted elaD showed weak DUB activity with K48-linked poly-ubiquitin by deconjugating Ub<sub>6</sub> to Ub<sub>4</sub>.

### **5.2.7 Analysis of elongation factors, both prokaryotic (EF-Tu) and eukaryotic (EF-1 $\alpha$ ) for DUB activity using unanchored K63/K48-linked mixed length poly-ubiquitin as substrates**

In Chapter 4 the candidates for the DUB activity in *E. coli* Rosetta<sup>TM</sup>2(DE3); the *E. coli* elongation factors EF-Tu (GST-TufA and GST-TufB) and GST-elaD, including the eukaryotic elongation factor (GST-EF-1 $\alpha$ 1 and GST-EF-1 $\alpha$ 2) were triaged using a ubiquitin-Sepharose pull down assay. This was following the hypothesis that a DUB activity would also be associated with the ability to bind ubiquitin as many DUBs contain ubiquitin binding domains (UBDs), some may not<sup>157</sup>. This theory was verified in this particular case, later through the use of gene knock-out techniques to knock-out elaD. It is possible that as well as elaD, there may be other DUBs in *E. coli* so it would be an interesting exercise to use poly-ubiquitin DUB assays to test TufA and TufB for any DUB activity.

A DUB assay was carried out using K63 and K48 linked poly-ubiquitin chain mixtures. GST (~15  $\mu$ g) and GST-tagged DUB candidate proteins (~15  $\mu$ g) were captured on glutathione-Sepharose 4B beads, then incubated (37°C) with 0.45  $\mu$ g either K63 (Fig.5.9.A) or K48 (Fig.5.9.B) linked poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose

membrane and probed for ubiquitin. Fig.5.11.A and Fig.5.11.B show GST-elaD was the only DUB candidate to deconjugate poly-ubiquitin chains (preferring K63-linked poly-ubiquitin). This confirms the choice of GST-elaD, during the triage of DUB candidates in Chapter 4.



**Fig. 5.11** GST and GST-tagged DUB candidate proteins were purified using glutathione-Sepharose 4B beads, and whilst still attached to glutathione-Sepharose 4B beads were then incubated (37°C) with either K63 (A) or K48 (B) linked poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. It is evident that GST-elaD was the only DUB candidate to demonstrate DUB activity as Ub<sub>5</sub> was deconjugated to Ub<sub>2</sub> although not completely as a faint band for Ub<sub>3</sub> remained. This verifies the choice of elaD as the likely DUB candidate chosen during the triage discussed in Chapter 4.

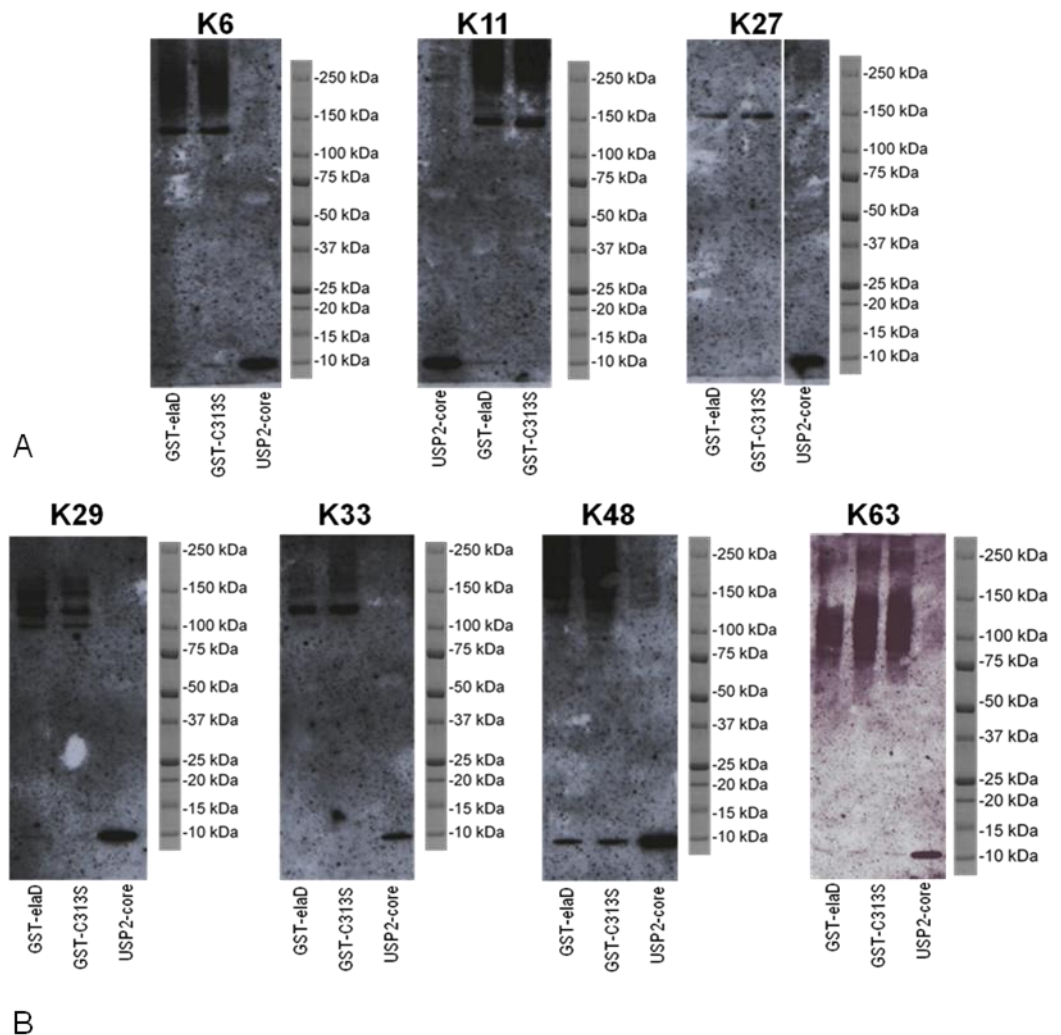


### **5.2.8 Anchored poly-ubiquitin chains as substrates for GST-elaD, including exclusive K6, K11, K27, K29, K33, K48, K63 linkages**

The poly-ubiquitin chains used as substrates for GST-elaD so far in this study have been unanchored K63-linked or K48-linked forms. Some eukaryotic DUBs, for example isopeptidase T (IsoT), have been known to demonstrate a preference for unanchored poly-ubiquitin chains over anchored poly-ubiquitin chains<sup>158</sup>. Therefore to investigate any preference for anchored poly-ubiquitin chains by GST-elaD, anchored chains were used. As only K63 and K48-linked unanchored poly-ubiquitin chains have been used previously in this study, all possible isomers of the anchored poly-ubiquitin chains were also included. These anchored poly-ubiquitin chains are GST-tagged, N-terminal to the polyubiquitin chain and were kindly supplied by BIOMOL- Enzo Life Sciences as they were not yet commercially available at the time that this work was carried out. These unanchored chains were first used by H. Wang *et al* in order to develop a specific antibody for K63-linked poly-ubiquitin<sup>159</sup>.

Eluted GST-tagged proteins (prepared as described in 5.2.6) were incubated (37°C) with 0.45 µg poly-ubiquitin chains (anchored) for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.5.12). It is

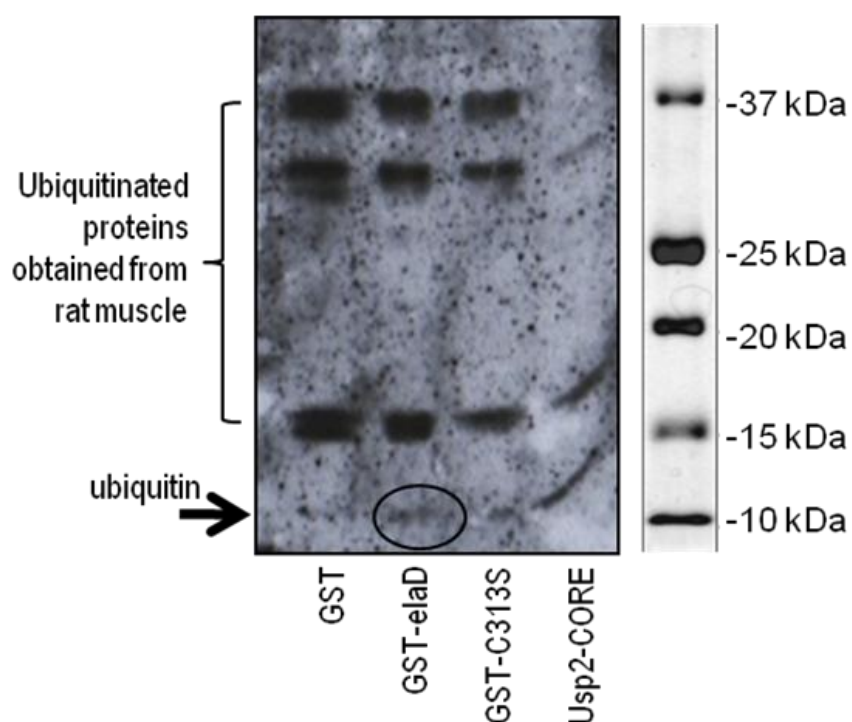
clear from Fig.5.12.B that GST-elaD did not deconjugate the anchored K63-linked poly-ubiquitin chains. Having demonstrated previously (5.2.4) that GST-elaD does deconjugate unanchored K63-linked poly-ubiquitin chains, this would indicate a preference by GST-elaD for unanchored poly-ubiquitin chains. GST-elaD did not deconjugate any of the other possible linkages of the unanchored poly-ubiquitin chains but the positive control did.



**Fig. 5.12** Anchored poly-ubiquitin chains containing one of either possible lysine-linkage: K6, K11, K27, K29, K33, K48 and K63, were used as a substrate for GST-elaD eluted from glutathione-Sepharose 4B beads. GST-elaD (5  $\mu$ g) and 5  $\mu$ g GST-tagged, catalytic mutant (C313S) were incubated (37°C) in parallel with 0.4  $\mu$ g USP2-core, a positive control DUB with 0.45  $\mu$ g poly-ubiquitin chain mixtures for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. GST-elaD did not deconjugate any of the forms of anchored poly-ubiquitin chains.

### **5.2.9 A mixture of target proteins, poly-ubiquitinated and mono-ubiquitinated, including unanchored poly-ubiquitin from rat muscle as substrates for GST-elaD**

First the ubiquitin binding protein, ZNF216 (expressed in *E. coli*) was attached to Sepharose 4B beads to generate ZNF216-Sepharose beads (2.4.2.4)<sup>160</sup>. Next, the hind leg muscle (35g) from a sacrificed adult male Lister-hooded rat (Charles River, UK) was homogenised in Tris-DTT (50mM Tris, 1mM DTT, pH7.5), including protease inhibitor cocktail on ice. Then, clarified rat muscle lysate containing a mixture of ubiquitinated proteins, was incubated with immobilised ZNF216-Sepharose beads (2.4.2.1). Eluted GST (~5 µg) or ~5 µg GST-tagged proteins (prepared as described in 5.2.6) were incubated (37°C) with immobilised rat muscle ubiquitinated proteins for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin (Fig.5.13). There is evidence in Fig.5.13 that GST-elaD did deubiquitinate some of the ubiquitinated proteins, as there is a faint band for mono-ubiquitin. Unanchored poly-ubiquitin was also likely to have been present as a substrate for elaD. It is possible that the weak band for mono-ubiquitin corresponds with the amount of unanchored chains rather than a weak DUB activity of GST-elaD.



**Fig. 5.13** A mixture of ubiquitinated proteins, captured from a rat muscle lysate, were used as substrates for GST-elaD. Eluted GST (5  $\mu$ g), GST-elaD (5  $\mu$ g) and the GST-tagged, 5  $\mu$ g catalytic mutant (C313S) were incubated (37°C) in parallel with USP2-core, a positive control DUB with immobilised ubiquitinated proteins on glutathione-Sepharose 4B beads for approximately 15 hr. The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. GST-elaD showed limited deubiquitinating activity as is evident from a faint band for mono-ubiquitin (circled).

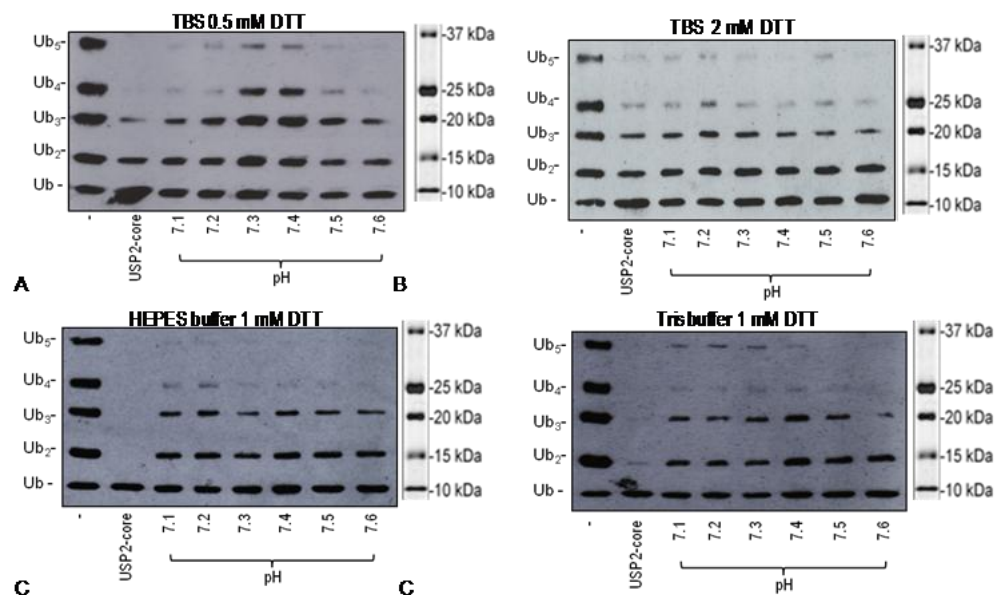
#### **5.2.10 Characterising GST-elaD DUB activity using unanchored K63-linked mixed length poly-ubiquitin**

The final analysis was to vary DUB assay conditions to investigate further the catalytic activity of GST-elaD, using K63-linked poly-ubiquitin as a preferred substrate. When carrying out assays involving DUB enzymes the reducing agent, dithiothreitol (DTT) was included to reduce disulphide bonds and therefore prevents unwanted intramolecular and intermolecular disulphide bonds from forming between cysteine residues. The concentration of DTT was varied (1 mM, 0.5 mM and 2 mM) to see how this affected the catalytic activity of GST-elaD.

Buffers differ in their ability to buffer pH, which in turn affects the optimum pH for the catalytic activity of an enzyme, therefore three buffers were used. The buffers used for the DUB assays were TBS (150 mM NaCl, 10 mM Tris), Tris buffer (50 mM Tris, 1 mM EDTA pH 7.5) and HEPES buffer (50 mM HEPES, 0.5 mM EDTA). The Tris buffer had been shown to provide optimal conditions for USP2-core, the positive control DUB used in this assay<sup>161</sup>. The HEPES buffer is a standard buffer used for DUB assays by ENZO Life Sciences UK (P. Sheppard, personal communication). Buffers were used in a range of pH values (7.1 – 7.6) measured at 37°C to find the optimum pH for the catalytic activity of GST-elaD. The assay was carried out using GST-elaD (15

μg) captured on Glutathione-Sepharose beads as described previously, then incubated with 0.45 μg unanchored K63-linked poly-ubiquitin and the proteins were analysed by probing for ubiquitin (Fig.5.14).

None of the assay conditions shown in Fig.5.14 were conducive to a complete deconjugation of K63-linked poly-ubiquitin chains by GST-elaD, showing that this enzyme has limited deubiquitinating activity. Previous DUB assays were carried out in TBS (0.5 mM DTT) at pH 7.5-7.6 (37 °C) which appear to be within the optimum range of pH for GST-elaD and as shown on previous occasions, for USP2-core activity for all buffers used in Fig.5.14. The USP2-core used in Fig.5.14.A and Fig.5.14.B were from an old stock so DUB activity was poor.



**Fig. 5.14** The DUB assay conditions were varied to investigate the ability of *elaD* to deconjugate K63 poly-ubiquitin chains over a range of pH 7.1 - 7.6 (37°C) and in different buffers. GST-*elaD* (15 µg) was purified using glutathione-Sepharose 4B beads, and whilst still attached to glutathione-Sepharose 4B beads, incubated (37°C) with 0.45 µg K63-linked poly-ubiquitin, 0.4 µg USP2-core was included as a positive control DUB. DUB assay samples were incubated (37°C) for 15 hr, for all 4 conditions, with TBS 0.5 mM DTT (A), TBS 2 mM DTT (B), HEPES buffer, 1 mM DTT (C) and Tris buffer, 1 mM DTT (D). The reaction was stopped using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, transferred on to nitrocellulose membrane and probed for ubiquitin. None of the assay conditions improved the ability of *elaD* to deconjugate chains shorter than Ub<sub>3</sub>, however *elaD* was more sensitive to pH when using TBS 0.5 mM DTT.



#### **5.2.11 Construction of mutations within GST-elaD designed to impair binding of GST-elaD to ubiquitin: GST-N227A, GST-W232A and GST-D169A**

Section 5.2.3 describes generating the GST-elaD catalytic mutant, GST-C313S, to compare with wild type GST-elaD in DUB assays and demonstrate that GST-elaD was responsible for the observed deconjugation of K63-linked poly-ubiquitin chains. The focus of this section now changes from the catalytic activity of GST-elaD to the binding of GST-elaD to ubiquitin, *i.e.* ubiquitin recognition by GST-elaD, previously observed in 4.2.4. This is of particular interest as this represents an interaction between a prokaryotic (elaD) and highly conserved eukaryotic (ubiquitin) protein.

A homology model of the structure of elaD was used to predict amino acid residues which may be responsible for catalysis and substrate binding. The homology model was based on the closest structural neighbour to elaD, *Xanthomonas* outer protein D (XopD), (personal communication by I. Dreveny, University of Nottingham) (Fig.5.15). XopD is a SUMO specific, deSUMOylating enzyme, a virulence factor secreted using Type III secretory system by *Xanthomonas campestris* *pv. Vesicatoria*. Results from homology modelling revealed six amino acid residues within elaD potentially important for binding or catalysis. Three of those amino acid residues (aspartate (D) 169, asparagine (N)

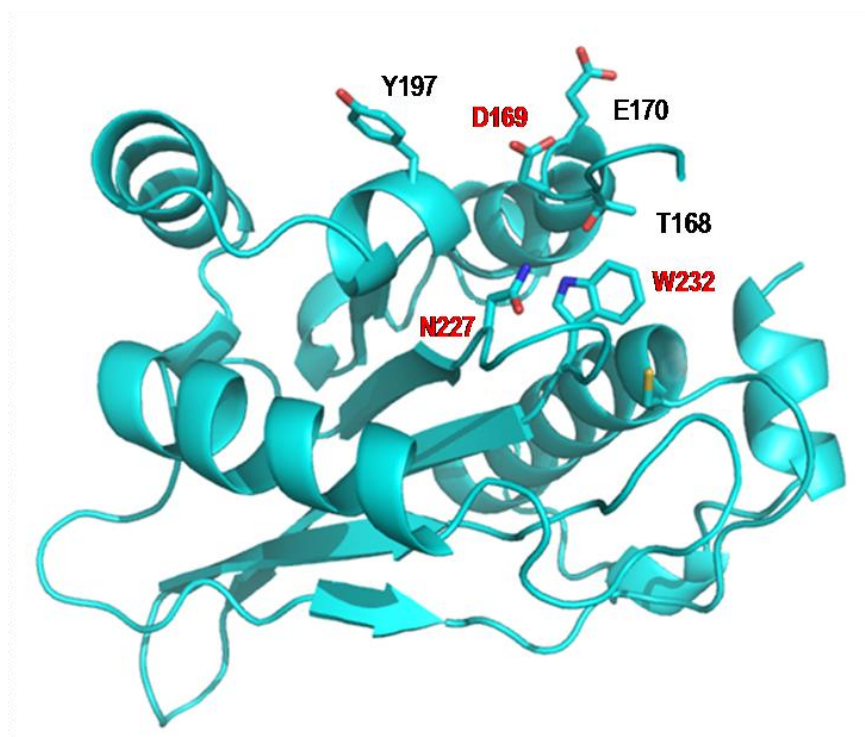
227 and tryptophan (W) 232) mutated in this study are shown in red in Fig.5.15 (the active site residues have been excluded).

The first of the three amino acid residues in *elaD*, chosen for mutation in an effort to disrupt binding or catalysis in this study was D169. D169 in *elaD* (conserved in XopD, all SENPs and SseL) and like NEDP1 could form a salt bridge with a conserved arginine residue at P35 position in its substrate. When the aspartate in NEDP1 (D29), corresponding with *elaD* D169, was mutated to D29A or D29N, NEDP1 activity was abolished. Therefore in this study, D169 in *elaD* was chosen for mutation to interfere *elaD* binding to ubiquitin<sup>117</sup>.

The second of the three amino acid residues in *elaD*, chosen for mutation to disrupt binding or catalysis, was N227. N227 in *elaD* (conserved in XopD, NEDP1, SENP6, SENP7, Ulp1 and SseL) may be important in forming a hydrogen bond with arginine in ubiquitin, as the conserved asparagine residue in NEDP1 (N91) can form a hydrogen bond with arginine residue 74, at the P3 position in NEDD8. The NEDP1 mutant, N91A was shown by L. Shen *et al* to be inactive, thus in this study N227 in *elaD* was chosen for mutation as it may potentially interfere with *elaD* binding to ubiquitin<sup>117</sup>.

The third of the three amino acid residues in *elaD*, chosen for mutation to disrupt binding or catalysis, was W232. W232 in *elaD* (conserved in XopD, all SENPs and SseL) can form Van der Waals interactions with glycine at P1 position in ubiquitin in the same way that NEDP1 does with NEDD8. A considerable loss of activity was reported by L. Shen *et al* when this conserved tryptophan in NEDP1 (W103) was mutated to W103A<sup>117</sup>. Therefore in this study, in an attempt to hinder *elaD* binding to ubiquitin, W232 was mutated.

Each of the three amino acid residues in *elaD*, D169, N227 and W232 were mutated to alanine, as the alanine side chain (-CH<sub>3</sub>) can not form a salt bridge, hydrogen bond or have Van der Waals interactions with ubiquitin. The three GST-*elaD* mutants, developed in an attempt to abrogate the ability of GST-*elaD* to bind ubiquitin were, GST-N227A, GST-W232A and GST-D169A. Wild type GST-*elaD* and the three mutants were then tested for their ability to bind ubiquitin using a ubiquitin-Sepharose pull-down assay. The three mutants, GST-N227S, GST-W232A and GST-D169A were generated by carrying out PCR based SDM using Pfu DNA polymerase as described in 2.6.4.



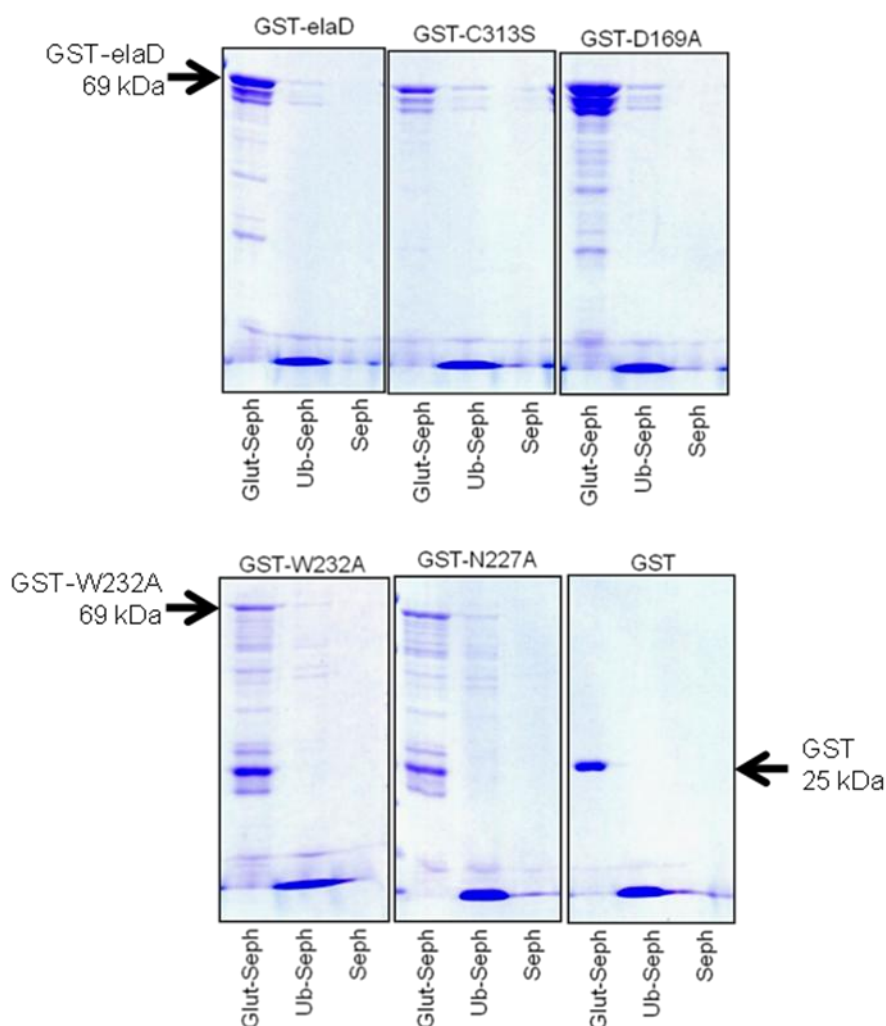
**Fig. 5.15** A homology model of the structure of the elaD peptidase domain based on the closest structural neighbour to elaD, *Xanthomonas* outer protein D (XopD), personal communication by I. Dreveny, University of Nottingham. XopD is specific for SUMO and is a virulence factor secreted using Type III secretory system by *Xanthomonas campestris* pv. *Vesicatoria*. The residues (excluding active site residues), are as indicated, and are predicted to be important for catalytic activity or substrate binding or for both functions. Mutants of elaD were generated by converting some of the residues (W232, D169 and N227) shown in red to alanine.

### **5.2.12 Using mutants of GST-elaD (N227A, W232A and D169A) to investigate recognition of ubiquitin**

Ubiquitin-Sepharose was used to test GST-tagged wild type (WT) elaD and elaD binding mutants, GST-N227A, GST-W232A and GST-D169A for their ability to bind ubiquitin. Protein was captured from a clarified lysate generated from a 10 ml culture, using a bead slurry (100  $\mu$ l) of either, glutathione-Sepharose 4B beads (glutathione-Sepharose), ubiquitin-Sepharose or Sepharose (control) and incubated (4°C) rotating for 1 hr, then washed with 0.1% (v/v) Triton X-100 in TBS. Proteins were eluted off the beads using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel and stained with Coomassie blue stain (Fig.5.16). When this assay was carried out previously the proteins were transferred on to a nitrocellulose membrane and probed for GST (4.2.4). However there was a large amount of non-specific protein which cross reacted with the GST antibody. Therefore to reduce sensitivity, this time a Coomassie stained gel was preferred.

When using Coomassie stain,  $\sim 1/10^{\text{th}}$  of wild type (WT) GST-elaD bound to glutathione-Sepharose was shown to bind to ubiquitin-Sepharose (Fig.5.16). This assumes that the glutathione-Sepharose reveals the total amount of GST-fusion available for binding. The predicted binding mutants, GST-W232A and GST-N227A showed reduced binding to ubiquitin-Sepharose, more so for GST-W232A

compared with WT GST-elaD. Interestingly, GST-D169A and to some extent GST-C313S appeared to show an increased ability to bind ubiquitin-Sepharose. The lack of binding by GST confirms the specificity of elaD for immobilised ubiquitin.



**Fig. 5.16** A ubiquitin-Sepharose pull-down assay to investigate any effects of mutating predicted ubiquitin binding residues may have over GST-elaD binding to ubiquitin. The three GST-tagged ubiquitin binding mutants tested were: D169A, W232A and N227A. The GST-elaD catalytic mutant, C313S was included to investigate any changes in binding ubiquitin as a result of changing a residue in the catalytic site. *E. coli* substrain XL10 Gold lysates were prepared in 0.1% (v/v) Triton X-100 in TBS as described 2.1.2. Protein was captured using either glutathione-Sepharose 4B beads (Glut-Seph), Ub-Seph or Seph and incubated (4°C) rotating for 1 hr, then washed with 0.1% (v/v) Triton X-100 in TBS. Proteins were eluted off the beads using gel loading buffer, electrophoresed on a 5-20% (w/v) acrylamide gradient gel, and Coomassie stained. W232A and N227A showed a reduced binding to ubiquitin, more so for W232A compared with WT. Interestingly, D169A and to some extent C313S showed an improved ability to bind ubiquitin.

### **5.3 Conclusions**

The work discussed in this chapter was aimed at developing a better understanding of the preference by GST-elaD for the topology of ubiquitin substrates and the mechanism of how GST-elaD binds to ubiquitin for catalysis. In this study (Chapter 3) it has already been demonstrated (*in vivo*) that GST-elaD is capable of cleaving linear peptide bonds in fusion proteins containing ubiquitin. The work in this chapter clearly shows that GST-elaD also has the ability to deconjugate isopeptide bonds within unanchored poly-ubiquitin chains and in particular, showed a preference for unanchored K63-linked poly-ubiquitin over K48-linked poly-ubiquitin chains.

The ability of GST-elaD to deconjugate K63-linked poly-ubiquitin is perhaps not surprising as the topology of linear ubiquitin is very similar to that of K63-linked poly-ubiquitin<sup>19</sup>. However, this preference for the K63-linked chains could indicate a possible function for GST-elaD. Eukaryotic DUBs targeting this poly-ubiquitin linkage are often involved in inhibiting the NF- $\kappa$ B signalling pathway, therefore preventing the expression of genes involved in the immune response<sup>6, 9, 15, 107, 108</sup>. Interestingly, both commensal and pathogenic *E. coli* (excluding extraintestinal strains) express elaD. Therefore these findings strengthen the argument that elaD is involved in some way in subverting host cell systems. However non-pathogenic *E. coli* lack the



specialised secretory structures found in pathogenic bacteria, for example the Type III secretory system which injects the host cell with virulence effectors. Non-pathogenic *E. coli* do have ABC transporters, as well as a general secretion (Sec) and two-arginine (Tat) translocation pathways, which could potentially be used for the export of *elaD*<sup>12, 162</sup>. Proteins secreted through the Sec and Tat pathways do need an N-terminal leader sequence for recognition by the systems. Assuming *elaD* has been translocated across the cytoplasmic membrane and cell wall of *E. coli* it would then need to be delivered in to the host cell. A novel form of contact-dependent signalling used by pathogens is also theoretically a possible way to transfer *elaD* directly into epithelial cell membranes<sup>163</sup>. This method would have to use novel transfer proteins as those known to be involved in this mode of transport have only been found in pathogenic bacteria<sup>164</sup>. More work needs to be done in that area.

The structure of *elaD* has yet to be determined, however fold recognition predictions of the peptidase domain show the closest structural neighbour to *elaD* to be XopD (*Xanthomonas* outer protein D) and SENP proteins, Ulp1 (yeast) and SENP2 (human) are also quite close. Therefore a paradox exists between the structure of *elaD* and its substrate, as the structural neighbours for *elaD* are deSUMOylating

enzymes and GST-elaD does not deSUMOylate but prefers ubiquitin. These findings were confirmed by A. Catic<sup>14</sup>.

There is evidence for a limited deubiquitinating activity with a mixture of ubiquitinated substrates from rat muscle (Fig.5.13), when used as a substrate for GST-elaD. However, this appears to be a weak DUB activity as it is less active than the positive control DUB, USP2-core. It is possible that the low level of mono-ubiquitin is related to the restriction of GST-elaD to unanchored K63-linked poly-ubiquitin than to a weak DUB activity in this case. When poly-ubiquitin was the substrate for GST-elaD only a weak DUB activity was observed. It is likely that there are some constraints over poly-ubiquitin recognition by GST-elaD, clear from the lack of deconjugation of poly-ubiquitin chains to below Ub<sub>2</sub> (Fig.5.10). However, there were still faint bands for Ub<sub>4</sub> and Ub<sub>5</sub>, which had not been deconjugated after 15 hr. The low level of catalytic activity by elaD in the current study was not observed by A. Catic *et al* when using ubiquitin-AMC as a substrate<sup>14</sup>. The low level activity of elaD seen in this study may be due to elaD having a higher affinity for ubiquitin-AMC than for the ubiquitin substrates used in the current study. Alternatively, one might suggest that this was an 'apparent' low level activity resulting from a low concentration of elaD in DUB assays. However if this was the case, one might expect that an active enzyme would have completed a reaction within a fifteen hour incubation time.

The most likely reason for the low level activity of GST-elaD observed in this study is due to the absence of a co-factor or another protein, necessary for GST-elaD to be fully active. GST-elaD was not a very active DUB when compared with the activities of other DUBs, reported by D. Komander *et al*<sup>19</sup>. In the study by Komander *et al*, nine DUBs, all selected from different groups of DUBs were incubated with either K48-linked, K63-linked or linear tetra-ubiquitin at a molar ratio (DUB : substrate), 4 : 1, for a range of times (0 min – 60 min). Isopeptidase T (IsoT) completely deconjugated K63-linked or K48-linked tetra-ubiquitin to within 10 minutes, compared with GST-elaD which had a molar ratio, 0.4 : 1 and incompletely deconjugated K63-linked poly-ubiquitin in 15 hr. The incomplete processing of poly-ubiquitin by GST-elaD is a feature of other enzymes for example, UCH-L1 and A20, had deconjugated tetra-ubiquitin to Ub<sub>3</sub> in an hour.

The weak binding of GST-elaD to ubiquitin (Fig.5.16) also shows that GST-elaD has a low affinity for mono-ubiquitin which probably contributes to the weak DUB activity. Limited information was obtained about the residues in GST-elaD important for binding to ubiquitin. Creating single point mutations in GST-elaD, predicted to perturb GST-elaD binding to immobilised ubiquitin revealed that W232 possibly contributes to GST-elaD binding to immobilised ubiquitin whereas N227 does not.

The absence of DUB activity by GST-elaD with some of the ubiquitin substrates was perhaps just as informative as the DUB activity observed. The lack of deconjugation of a linear peptide bond in di-ubiquitin was initially unexpected as GST-elaD has been shown to cleave the linear peptide bond in fusion proteins containing ubiquitin. However, GST-elaD did only deconjugate unanchored K63-linked poly-ubiquitin to tri-ubiquitin, so perhaps GST-elaD lacks necessary recognition sites for binding to di-ubiquitin, regardless of the bond type between the ubiquitin molecules. It is particularly interesting that GST-elaD did not deconjugate anchored K63-linked poly-ubiquitin chains yet did deconjugate unanchored poly-ubiquitin of the same isomer. This preference for unanchored poly-ubiquitin is associated with IsoT, a DUB which deconjugates ubiquitin chains in the cell to recycle ubiquitin<sup>165</sup>. IsoT requires the C-terminus of the proximal ubiquitin in the poly-ubiquitin chain to be free for binding. These findings lead to the suggestion that when poly-ubiquitin is the substrate, GST-elaD needs the C-terminus of the proximal ubiquitin in the chain in order to bind and deconjugate the chain. A slightly different mechanism must operate when GST-elaD deubiquitinates a linear fusion protein as ubiquitin is connected to a fusion partner and therefore there is no exposed ubiquitin C-terminus for recognition.

From the work discussed in this chapter, more has been learned about the preference GST-elaD has for ubiquitin isomers as substrates. GST-elaD prefers unanchored chains, and specifically prefers K63-linked poly-ubiquitin. elaD is only expressed in commensal and pathogenic *E. coli* associated with the intestines. This and the preference for K63-linked poly-ubiquitin could indicate that elaD plays a role in modulating the inflammatory response in the intestines to retain a symbiotic relationship with host cells. More work needs to be carried out to investigate if this is the case and to ascertain how this occurs.

## 6 Discussion

### 6.1. Key findings of this study

#### 6.1.1 *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) expresses a specific DUB with a conventional mechanism

The work in this study has clearly demonstrated that *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) has a specific, intrinsic DUB activity. Work leading up to this project involved the observation of a novel ubiquitin-fusion processing activity in a DE3 strain of *E. coli* by other members of this group (H. Jefferey and I. Kerr, unpublished). The processing activity occurred *in vivo*, against an N-terminal hexahistidine tagged ubiquitin-lincomycin resistance C fusion protein, (His)<sub>6</sub>-Ub-LmrC expressed in the DE3 strain of *E. coli*. The initial work in this study was aimed at characterising the novel ubiquitin-fusion processing activity, initially in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). The DUB activity was specific (i.e. cleaved post Gly76 of ubiquitin), confirmed by N-terminal sequence analysis (Chapter 3). A main focus of this study was to investigate the specificity of the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) as well as to identify this prokaryotic DUB. The initial work showed that the DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) had a conventional catalytic activity. The use of (His)<sub>6</sub>-Ub-LmrC mutants with point mutations around the DUB cleavage site (GA, GP and GGP), designed to perturb

cleavage showed that the prokaryotic DUB, like most mammalian DUBs (D. A. Gray, personal communication) did not cleave bonds when proline forms part of the DUB cleavage site. The inability of the prokaryotic DUB to cleave GA (i.e. G76A mutation at the C-terminus of ubiquitin) was similar to the findings of Butt and co-workers who previously characterised the specificity of DUBs expressed in yeast<sup>148</sup>. In this study it was demonstrated that GST-elaD has a conventional catalytic mechanism, one which requires cysteine within its catalytic triad. This is shown by the lack of activity by the GST-elaD catalytic mutant C313S compared with the activity of wild type GST-elaD in DUB assays (*in vitro*). The lack of catalytic activity by C313S was not simply the result of poor binding to substrate, as C313S was better at binding with ubiquitin (in ubiquitin-Sepharose) than wild type GST-elaD.

#### **6.1.2 The DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was shown to prefer ubiquitin over SUMO1 or NEDD8 as substrates *in vivo***

The work in this study was aimed at investigating any preference that the DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) may have for ubiquitin or the ubiquitin-like proteins (UBLs) small ubiquitin-related modifier 1 (SUMO1) or neural precursor cell expressed developmentally down-regulated protein 8 (NEDD8). If a preference was shown by the DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) then this could give more information

about its catalytic site/mechanism in comparison with known DUBs or ubiquitin-like proteases (ULPs) and may even indicate a possible function. Some DUBs or ULPs are exclusive to their substrates, whilst others are less specific. For example the deNEDDylating ULP, NEDP1 (human) is specific for NEDD8 only, whereas the DUB, USP21 (human) can cleave ubiquitin and NEDD8 but not SUMO1<sup>109, 166</sup>. The deSUMOylating ULPs (all human), SENP1, SENP2 and SENP3 are all specific for SUMO. The specificity of DUBs or ULPs for their substrates is still not fully understood, however, functional studies and structural information have helped us to understand some structural controls over specificity.

Shen *et al.* determined the crystal structures of NEDP1 bound to NEDD8 or in isolation and compared those findings with the structure for, Yuh1 (yeast) covalently linked to ubiquitin aldehyde determined by Johnston *et al.* and the structure for Ulp1 (yeast) in complex with SUMO determined by Mossessova *et al.*<sup>117, 167, 168</sup>. Shen *et al.* highlighted the importance of the 72<sup>nd</sup> amino acid residue in NEDD8 and ubiquitin in determining the specificity of NEDP1 for NEDD8 over ubiquitin. The 72<sup>nd</sup> amino acid residue of ubiquitin is arginine (R) Ub<sup>R72</sup> and NEDD8 is alanine (A) NEDD8<sup>A72</sup>. The specificities of NEDP1 and Ulp1 for their substrates, NEDD8 and SUMO1, respectively, are explained by Shen *et al.* as being due to two structural characteristics of the proteases. The



surfaces on each protease have very different charges, which ensure specificity for the substrate. Additionally, each protease has a loop which moves over the C-terminus of each protease in complex with the substrate and in NEDP1 this loop is much longer than in Ulp1 limiting binding partners for each protease and ensuring strict specificity. The findings of our study have demonstrated that the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) preferred ubiquitin over SUMO1 or NEDD8. These initial investigations compared the ability of fusion proteins containing either ubiquitin, NEDD8 or SUMO1 to be cleaved, *in vivo* when expressed in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3). Only the fusion protein containing ubiquitin was cleaved *in vivo*. Therefore, this would indicate that the DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) has a catalytic site and a binding surface more comparable to Yuh1. The finding in our study that the DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) shows a preference for ubiquitin over NEDD8 or SUMO1 is significant. *E. coli* do not have a system like the eukaryotic ubiquitin-mediated system and therefore have no ubiquitin homologue to function as a protein tag. It is possible that the two ubiquitin homologues, ThiS and MoaD which take part in the biosynthesis of thiamin and Molybdenum cofactor (respectively) in *E. coli* are substrates for elaD. The possible implications for a prokaryotic DUB having a specificity for ubiquitin will be discussed later in section 6.1.6.

### **6.1.3 The DUB in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was identified as elaD**

At the start of this study there were no reports of DUB expression in *E. coli*. It was known however from the literature that some prokaryotic DUBs were expressed by pathogenic eubacteria other than *E. coli*, and that these DUBs were involved in subversion of host cell systems for virulence<sup>6-8, 15, 107 9, 10, 108</sup>. During the progression of this study a ULP, elaD, expressed by *E. coli*, was discovered by Catic in a group lead by Ploegh<sup>14</sup>. Using a bioinformatics search for new members of a clan containing UBPs and ULPs Catic *et al.* identified elaD and subsequently showed that GST-elaD had DUB activity, using suicide inhibitors and ubiquitin-AMC as substrates for elaD. Therefore the work in our study aimed at identifying the identity of the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) included elaD as a candidate. The prokaryotic elongation factor, EF-Tu, shown by Ciechanover and colleagues to be associated with an ubiquitin protease activity, was also included as a candidate for the activity<sup>139</sup>. The identity of the DUB activity in *E. coli* substrain Rosetta<sup>TM</sup>2(DE3) was established as elaD, using a gene knock-out technique to knock-out elaD, creating Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>apr</sup>. When fusion proteins shown to cleave in the parental *E. coli*, (His)-Ub-LmrC or (His)<sub>6</sub>-Ub-enhanced green

fluorescence protein (EGFP), were expressed in Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>apr</sup>, DUB activity was abolished (4.2.3).

The findings in our study (6.1.2) of the preference for ubiquitin over NEDD8 or SUMO1 by the prokaryotic DUB, now known to be elaD, are supported by the biochemical tests carried out by Catic *et al.*<sup>14</sup>. The authors found that GST-elaD cleaved ubiquitin-AMC but not SUMO-AMC or NEDD8-AMC. However, the use of suicide inhibitors showed that GST-elaD became covalently attached to ubiquitin-vinyl methylester but not SUMO-vinylsulphone (SUMOVS) and to a limited extent GST-elaD also became covalently linked to NEDD8VS. The findings by Catic *et al.* suggest a weak DUB activity by elaD for NEDD8. In our study we observed no cleavage of NEDD8 in (His)<sub>6</sub>-NEDD8-LmrC; perhaps these different findings relate to the differences in NEDD8 substrates. It is possible that the binding of GST-elaD to NEDD8 is an artefact due to the use of suicide inhibitors. The NEDD8VS suicide inhibitor is slightly different from NEDD8. The suicide inhibitors developed by Catic *et al.* have an alternative electrophilic region neighbouring the electrophilic carbonyl carbon (which normally undergoes nucleophilic attack by the DUB). Additionally, the electrophilic region has to be stabilised by an electron withdrawing group, therefore the chemical environment in NEDD8 is different from that of NEDD8VS. The positioning of NEDP1 by

NEDD8<sup>A72</sup> was found to be a critical factor in conferring specificity to NEDP1 for NEDD8<sup>117</sup>. Therefore, the slight difference in peptide bond location and in the chemical environments of NEDD8VS compared with NEDD8 may be enough of a change to favour a binding that would not normally occur with NEDD8.

#### **6.1.4 GST-elaD prefers unanchored Lysine 63-linked poly-ubiquitin as a substrate *in vitro***

In this study work carried out (discussed in Chapter 5) using different types of ubiquitin substrates for GST-elaD was aimed at building a picture of the specificity that GST-elaD has for ubiquitin. In a recent review, Komander *et al.* describe the five different types of specificities displayed by DUBs for poly-ubiquitin chains<sup>130</sup>. DUBs can be linkage-specific, for example to prefer Lysine 48-linked (K48) poly-ubiquitin; regionally-specific, for example either *endo*-specific or *exo*-specific; substrate-specific; mono-ubiquitin-specific or specific for unanchored ubiquitin.

The findings of our study have demonstrated that GST-elaD showed a clear preference for Lysine 63-linked (K63) poly-ubiquitin chains over K48-linked poly-ubiquitin chains. Interestingly, as well as being linkage specific for K63-linked poly-ubiquitin, GST-elaD also demonstrated a specificity for unanchored poly-ubiquitin as GST-elaD was unable to

deconjugate anchored K63-linked poly-ubiquitin (5.2.8). This could mean that GST-elaD is either specific for unanchored chains, or GST-elaD is substrate-specific with the added ability in being able to deconjugate unanchored K63-linked poly-ubiquitin. If GST-elaD simply has a preference for unanchored poly-ubiquitin this would suggest that, with poly-ubiquitin elaD must have a ubiquitin recognition site similar to that in the eukaryotic DUB, isopeptidase T (IsoT)<sup>158, 165</sup>. The lack of specificity of IsoT for unanchored poly-ubiquitin, i.e. it is not linkage-specific, makes it ideal for recycling K63-linked poly-ubiquitin (and possibly linear poly-ubiquitin) from poly-ubiquitin chains. However, if elaD is substrate-specific it is possible that elaD deubiquitinates a protein in a eukaryotic signalling pathway.

The differences between the observations in this study of GST-elaD DUB activity and what is known about catalytic mechanism of IsoT, leads to the conclusion that, if elaD is specific for unanchored poly-ubiquitin then elaD has a slightly different DUB mechanism to IsoT. GST-elaD was able to cleave ubiquitin fusion proteins, (His)<sub>6</sub>-Ub-LmrC and (His)<sub>6</sub>-Ub-EGFP (*in vivo*). There is no exposed GG of the C-terminus of ubiquitin in either of the fusion proteins for GST-elaD to recognise, so this leads one to suggest that elaD has an alternative site for the recognition of ubiquitin. The cleavage of the fusion proteins in this study (*in vivo*) may be restricted to fusion proteins as nascent poly-

peptides, so in an unfolded state, therefore revealing extra recognition sites necessary for elaD. However, elaD could be substrate-specific and therefore does not have the strict requirement for a C-terminus on ubiquitin.

The findings of our study also showed that GST-elaD does not appear to recognise di-ubiquitin as GST-elaD did not deconjugate linear di-ubiquitin and only deconjugated unanchored K63-linked poly-ubiquitin to di-ubiquitin. Although the bonds between ubiquitin molecules are different in linear di-ubiquitin compared with those in K63-linked poly-ubiquitin, the peptide bond and K63-linked isopeptide bonds are structurally similar<sup>19</sup>. Therefore it is possible that GST-elaD may deconjugate linear poly-ubiquitin chains longer than n=2. However, this was not tested in this study. Constraints over time prevented the generation of linear ubiquitin substrates using an *in vitro* system that included the E3, the linear ubiquitin assembly complex (LUBAC) and linear poly-ubiquitin is not yet commercially available.

#### **6.1.5 GST-elaD-ubiquitin binding and elaD structural neighbours**

Limited information was obtained about the amino acid residues responsible for binding of GST-elaD to ubiquitin. GST-elaD mutants N227A, W232A and D169A were generated to perturb the ability of

GST-elaD to bind ubiquitin. A ubiquitin-Sepharose pull-down was used to compare the abilities of GST-tagged WT and mutant elaD to bind immobilised ubiquitin. Only residues W232 and N227 were demonstrated to be important for binding, but not vital as limited binding still occurred. More work needs to be carried out in this way to test other amino acid residues predicted to be important for binding elaD to ubiquitin.

As part of this work a homology model of the structure of elaD was drawn based upon *Xanthomonas* outer protein D (XopD), the closest structural neighbour to elaD, (personal communication by I. Dreveny, University of Nottingham). Interestingly, XopD is a deSUMOylating enzyme which is specific for SUMO. This contrasts with elaD as although it is structurally similar to XopD, elaD does not cleave SUMO1 and is specific for ubiquitin. This mismatch between structure and function of proteases is not unusual, for example, Avp is an adenoviral protease, structurally similar to the deSUMOylating protease, ULP1, yet Avp does not cleave SUMO, it cleaves a ubiquitin homologue precursor protein, ISG15<sup>169</sup>.

#### **6.1.6 Implications and outcomes**

The preference of elaD for ubiquitin, a eukaryotic protein, is significant, as it indicates that elaD may have a function associated with ubiquitin in

eukaryotic cells. *E. coli* do not have a system like the eukaryotic ubiquitin-mediated systems, and do not express ubiquitin or a ubiquitin homologue used as a protein tag. Therefore it is unlikely that this preference by GST-elaD for ubiquitin is related to a function within bacteria. Under standard growth conditions elaD did not appear to be essential for *E. coli* survival, as there was no evidence of impaired growth of the *E. coli* strain with elaD knocked out (Rosetta<sup>TM</sup>2(DE3) $\Delta$ elaD<sup>apr</sup>), compared with wild type *E. coli* Rosetta<sup>TM</sup>2(DE3).

This, together with the observation that elaD is only expressed in commensal *E. coli* or pathogenic *E. coli* associated only with the intestines are indicators of elaD having a function which benefits the existence of a commensal bacterium. This could be to evade an immune response by dampening host cell signalling systems<sup>14</sup>. If this is the case, elaD could potentially benefit both commensal and pathogenic intestinal *E. coli*, but a secretion system and delivery mechanism would be needed.

Commensal *E. coli* do have secretion systems and they certainly have been shown to dampen NF- $\kappa$ B signalling in intestinal epithelial cells, dendritic cells and B cells. The precise details of how these effects are achieved, however, are not known. The commensal bacterium,



*Bacteriodes thetaiotamicron* was shown to induce the export of RelA, one of the subunits of the transcription factor, NF- $\kappa$ B from the nucleus, preventing transcription of inflammatory genes<sup>170</sup>. *B. thetaiotamicron* was grown in a co-culture with an immortalised intestinal epithelial cell line, human epithelial colorectal adenocarcinoma (Caco-2) cells. It was demonstrated that *B. thetaiotamicron* induced the transcription of PPAR- $\gamma$  and the movement of peroxisome proliferator activated receptor- $\gamma$  (PPAR- $\gamma$ ) in to the nucleus of Caco-2 cells followed by their exit from the nucleus with RelA. Another commensal bacterium, *Lactobaccillus casei* has been reported to prevent proteolysis of I $\kappa$ B by the proteasome by an unidentified mechanism<sup>171</sup>.

These examples demonstrate that commensal bacteria are capable of secreting effectors to modulate host cell systems, though more work is needed to identify the bacterial effectors and how they are delivered to the host. If *elaD* has a dampening effect over NF- $\kappa$ B signalling in host cells more needs to be learned of how commensal bacteria achieve this to understand any role *elaD* may have in *E. coli*.

In contrast to the limited knowledge of how commensal bacteria dampen NF- $\kappa$ B signalling in host cells, more is known of how pathogenic bacteria of the gut achieve this. However, if *elaD* is simply

specific for unanchored K63-linked poly-ubiquitin it is unlikely that elaD would use the same mechanisms as known pathogens.

*Yersinia pseudotuberculosis* (described in Chapter 1) expresses a DUB, YopJ which deubiquitinates K63-linked poly-ubiquitin chains from TRAF2 and TRAF6 and K48-linked poly-ubiquitin chain from I $\kappa$ B therefore inhibiting production of inflammatory proteins<sup>15</sup>. Therefore YopJ is a DUB with dual specificity for K63-linked and K48-linked poly-ubiquitin isomers and cleaves anchored chains. The findings of our study showed that elaD prefers unanchored K63-linked poly-ubiquitin over anchored K63-linked poly-ubiquitin. Therefore it would be unlikely that elaD could deubiquitinate TRAF 2, TRAF 6 or I $\kappa$ B as the chains would be anchored. The preference for unanchored poly-ubiquitin by elaD would suggest that elaD has a similar role to that of IsoT, which recycles ubiquitin from poly-ubiquitin chains<sup>158</sup>. elaD could recycle K63-linked poly-ubiquitin and linear unanchored chains. The lack of availability of linear chains longer than di-ubiquitin has precluded the use of linear poly-ubiquitin as a substrate for elaD in this study. It is likely that elaD could also process ubiquitin precursor proteins as GST-elaD has been demonstrated in this study to cleave ubiquitin fusion proteins. IsoT does have a preference for K63-linked and K48-linked unanchored poly-ubiquitin over unanchored linear poly-ubiquitin, so elaD could potentially improve the recycling of unanchored K63-linked

and unanchored linear poly-ubiquitin ubiquitin<sup>19</sup>. However, this potential improvement in the availability of ubiquitin by elaD is unlikely to have any real benefit in dampening NF- $\kappa$ B signalling.

Alternatively, if elaD is substrate-specific, rather than simply specific for unanchored K63-linked poly-ubiquitin, it is feasible that elaD could target NEMO, TRAF2, TRAF3 or TRAF6, or all three TRAF proteins depending on the degree of specificity that elaD may have for the substrate. NEMO has been shown to be ubiquitinated by an E3, the linear ubiquitin assembly complex (LUBAC) which builds linear poly-ubiquitin chains with the same effects as K63-linked poly-ubiquitin, recently reviewed<sup>172</sup>. elaD could also potentially deconjugate linear poly-ubiquitin in this context.

It is intriguing that a DUB expressed in both enteropathogenic *E. coli* and commensal *E. coli*, has an affinity for eukaryotic ubiquitin and neither *E. coli* strain expresses a ubiquitin homologue. elaD could be a key enzyme in modulating host cell systems and further work promises to improve our understanding of the relationship between commensal and enteropathogenic *E. coli* and host cells.

## **6.2. Future work**

More work needs to be done to explore the possible role for elaD and whether this is beneficial to commensal *E. coli*, enteropathogenic *E. coli* or both. The catalytic mechanism of elaD needs further investigation, as well as to search for possible binding partners for elaD.

### **6.2.1 Investigating the possible role or roles of elaD**

An investigation could be carried out to test the hypothesis that elaD may be important for dampening NF- $\kappa$ B signalling in both enteropathogenic and commensal *E. coli*. Therefore elaD could be knocked out of an enteropathogenic strain of *E. coli*, for example O157 and this strain could be compared with parental O157 in their abilities to invade eukaryotic cells. The work could include assays used to measure changes in localised actin polymerisation within eukaryotic cells and the appearance of pedestal structures in epithelial tissue.

To discover how elaD may be delivered to host cells an antibody against elaD and fluorescent microscopy could be used to detect any secretion of elaD. Assays could also be developed to test for contact-dependent transfer of elaD.

The importance of individual domains in *elaD* in determining pathogenicity of *E. coli* could be investigated. *elaD* is thought to have two domains (K. Hoffman, personal communication), a ULP-like domain, ~128<sup>th</sup> – 407<sup>th</sup> residue of *elaD* and the first 120 residues probably form a separate domain which could be important for the catalytic domain or may be important for recognising interaction partners or associating with type III secretion systems. It would be interesting to test the hypothesis that the first 120 residues of *elaD* might play a key role in determining the pathogenic state of *E. coli*. It may be possible that in enteropathogenic *E. coli* the first 120 residues of *elaD* carry a mutation or are truncated, favouring an interaction with a different protein and enabling virulence. A bioinformatics study could be carried out, starting with a BLAST search for *elaD* and comparing the first 120 residues. Immobilised wild type *elaD* could be used in a pull-down from eukaryotic lysate and *E. coli* lysate to search for interacting proteins. If this is successful, constructs of *elaD*, from various pathogenic *E. coli* could be generated and used in similar pull-downs to see if this changes binding partners.

*elaD* could be more important for commensal *E. coli* than for enteropathogenic *E. coli* which would have the added advantage of many virulence effectors and secretory systems. This would require initial work looking in to the effects of wild type commensal *E. coli* over,

NF- $\kappa$ B signalling, PPAR- $\gamma$  induction and RelA export from the nucleus, in intestinal epithelial cells. This could then be followed by similar work comparing the effects of *E. coli* with *elaD* knocked out with WT commensal *E. coli*.

### **6.2.2 Further investigations in to the catalytic activity and interaction partners for *elaD***

When available, unanchored linear poly-ubiquitin could be used as a substrate for *elaD* to see if there is a greater preference for unanchored linear chains over unanchored K63-linked poly-ubiquitin. To investigate if *elaD* is substrate-specific, *elaD*-Sepharose could be generated and used to search for substrates in eukaryotic lysates. NEMO, TRAF2, TRAF3 and TRAF6 could be ubiquitinated with K63-linked poly-ubiquitin (including linear poly-ubiquitinated NEMO), *in vitro* and used as substrates for *elaD*. This use of immobilised *elaD* could also potentially identify any other interaction partners for *elaD* in eukaryotic and in *E. coli* lysates.

*elaD* could be truncated or mutated and used in DUB assays to see if the first 120 residues of *elaD* are important for catalysis. If wild type immobilised *elaD* succeeded in pulling down interaction proteins then immobilised, truncated or mutated *elaD* could be used in a pull-down compared with immobilised wild type *elaD*. DUB assays using

pathogenic *elaD* constructs could be compared with wild type *elaD* to test for different DUB activities.

Investigation in to the possible effects that pH may have over specificity of *elaD* for K48-linked poly-ubiquitin may also be relevant. K48-linked poly-ubiquitin adopts a fully closed conformation at pH 7.5 and is fully open at pH 4.5<sup>16</sup>. The DUB assays carried out in this study were at pH 7.6 (37°C), therefore the K48-linked poly-ubiquitin chains would have been in a closed conformation. A report by Karagiannis and Young found the pH of yeast to be maintained at pH 7.3 throughout the cycle. However minor stresses can lower pH to 5.75<sup>173</sup>. If these values are extrapolated to intestinal epithelial cells then K48-linked poly-ubiquitin would be partially open and more so in cases of cell stress, due to for example, a challenge by pathogenic bacteria. Therefore, it is possible that, some DUBs which appear to be linkage-specific *in vitro* may not be linkage-specific *in vivo*.

## 7 References

1. Erbse, A. *et al.* ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. *Nature* **439**, 753-756 (2006).
2. Graciet, E. *et al.* Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *Proc. Natl. Acad. Sci. U. S. A* **103**, 3078-3083 (2006).
3. Jonnalagadda, S. *et al.* Expression and accurate processing of yeast penta-ubiquitin in *Escherichia coli*. *J. Biol. Chem.* **262**, 17750-17756 (1987).
4. Catanzariti, A.M., Soboleva, T.A., Jans, D.A., Board, P.G., & Baker, R.T. An efficient system for high-level expression and easy purification of authentic recombinant proteins. *Protein Sci.* **13**, 1331-1339 (2004).
5. Varshavsky, A. The N-end rule pathway of protein degradation. *Genes Cells* **2**, 13-28 (1997).
6. Le, N.G. *et al.* ChlA<sub>Dub1</sub> of *Chlamydia trachomatis* suppresses NF- $\kappa$ B activation and inhibits I $\kappa$ B $\alpha$  ubiquitination and degradation. *Cell Microbiol.* **10**, 1879-1892 (2008).
7. Ye, Z., Petrof, E.O., Boone, D., Claud, E.C., & Sun, J. Salmonella effector AvrA regulation of colonic epithelial cell inflammation by deubiquitination. *Am. J. Pathol.* **171**, 882-892 (2007).
8. Rytönen, A. *et al.* SseL, a Salmonella deubiquitinase required for macrophage killing and virulence. *Proc. Natl. Acad. Sci. U. S. A* **104**, 3502-3507 (2007).



9. Haase,R., Richter,K., Pfaffinger,G., Courtois,G., & Ruckdeschel,K. Yersinia outer protein P suppresses TGF-beta-activated kinase-1 activity to impair innate immune signaling in Yersinia enterocolitica-infected cells. *J. Immunol.* **175**, 8209-8217 (2005).
10. Thiefes,A. *et al.* The Yersinia enterocolitica effector YopP inhibits host cell signalling by inactivating the protein kinase TAK1 in the IL-1 signalling pathway. *EMBO Rep.* **7**, 838-844 (2006).
11. Chen,Z.J. & Sun,L.J. Nonproteolytic functions of ubiquitin in cell signaling. *Mol. Cell* **33**, 275-286 (2009).
12. Tseng,T.T., Tyler,B.M., & Setubal,J.C. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC. Microbiol.* **9 Suppl 1**, S2 (2009).
13. Yang,Y., Li,C.C., & Weissman,A.M. Regulating the p53 system through ubiquitination. *Oncogene* **23**, 2096-2106 (2004).
14. Catic,A., Misaghi,S., Korbel,G.A., & Ploegh,H.L. ElaD, a Deubiquitinating protease expressed by E. coli. *PLoS. ONE.* **2**, e381 (2007).
15. Zhou,H. *et al.* Yersinia virulence factor YopJ acts as a deubiquitinase to inhibit NF-kappa B activation. *J Exp. Med.* **202**, 1327-1332 (2005).
16. Varadan,R., Walker,O., Pickart,C., & Fushman,D. Structural properties of polyubiquitin chains in solution. *J. Mol. Biol.* **324**, 637-647 (2002).
17. Erez,E., Fass,D., & Bibi,E. How intramembrane proteases bury hydrolytic reactions in the membrane. *Nature* **459**, 371-378 (2009).

18. Xirodimas,D.P., Saville,M.K., Bourdon,J.C., Hay,R.T., & Lane,D.P. Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* **118**, 83-97 (2004).
19. Komander,D. *et al.* Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep.* **10**, 466-473 (2009).
20. Fujinaga,M., Cherney,M.M., Oyama,H., Oda,K., & James,M.N. The molecular structure and catalytic mechanism of a novel carboxyl peptidase from *Scytalidium lignicolum*. *Proc. Natl. Acad. Sci. U. S. A* **101**, 3364-3369 (2004).
21. Orłowski,M. & Wilk,S. Catalytic activities of the 20 S proteasome, a multicatalytic proteinase complex. *Arch. Biochem. Biophys.* **383**, 1-16 (2000).
22. Gostissa,M. *et al.* Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *EMBO J.* **18**, 6462-6471 (1999).
23. Hershko,A. & Ciechanover,A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425-479 (1998).
24. Jentsch,S. The ubiquitin-conjugation system. *Annu. Rev. Genet.* **26**, 179-207 (1992).
25. Jung,T., Catalgol,B., & Grune,T. The proteasomal system. *Mol. Aspects Med.* **30**, 191-296 (2009).
26. Wolf,S., Lottspeich,F., & Baumeister,W. Ubiquitin found in the archaebacterium *Thermoplasma acidophilum*. *FEBS Lett.* **326**, 42-44 (1993).
27. Jap,B. *et al.* Preliminary X-ray crystallographic study of the proteasome from *Thermoplasma acidophilum*. *J. Mol. Biol.* **234**, 881-884 (1993).

28. Durner, J. & Boger, P. Ubiquitin in the prokaryote *Anabaena variabilis*. *J Biol. Chem.* **270**, 3720-3725 (1995).
29. Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., & Darwin, K.H. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* **322**, 1104-1107 (2008).
30. Darwin, K.H. Prokaryotic ubiquitin-like protein (Pup), proteasomes and pathogenesis. *Nat. Rev. Microbiol.* **7**, 485-491 (2009).
31. DeMartino, G.N. PUPylation: something old, something new, something borrowed, something Glu. *Trends Biochem. Sci.* **34**, 155-158 (2009).
32. Iyer, L.M., Burroughs, A.M., & Aravind, L. Unraveling the biochemistry and provenance of pupylation: a prokaryotic analog of ubiquitination. *Biol. Direct.* **3**, 45 (2008).
33. Kiel, C. & Serrano, L. The ubiquitin domain superfold: structure-based sequence alignments and characterization of binding epitopes. *J Mol. Biol.* **355**, 821-844 (2006).
34. Gill, G. SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev.* **18**, 2046-2059 (2004).
35. Cardone, L. *et al.* Circadian clock control by SUMOylation of BMAL1. *Science* **309**, 1390-1394 (2005).
36. Giannakopoulos, N.V. *et al.* Proteomic identification of proteins conjugated to ISG15 in mouse and human cells. *Biochem. Biophys. Res. Commun.* **336**, 496-506 (2005).

37. Zhao,C., Denison,C., Huibregtse,J.M., Gygi,S., & Krug,R.M. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc. Natl. Acad. Sci. U. S. A* **102**, 10200-10205 (2005).
38. Wu,J.T., Lin,H.C., Hu,Y.C., & Chien,C.T. Neddylation and deneddylation regulate Cul1 and Cul3 protein accumulation. *Nat. Cell Biol.* **7**, 1014-1020 (2005).
39. Ohsumi,Y. Molecular dissection of autophagy: two ubiquitin-like systems. *Nat. Rev. Mol. Cell Biol.* **2**, 211-216 (2001).
40. Thompson,A.R., Doelling,J.H., Suttangkakul,A., & Vierstra,R.D. Autophagic nutrient recycling in Arabidopsis directed by the ATG8 and ATG12 conjugation pathways. *Plant Physiol* **138**, 2097-2110 (2005).
41. Goehring,A.S., Rivers,D.M., & Sprague,G.F., Jr. Attachment of the ubiquitin-related protein Urm1p to the antioxidant protein Ahp1p. *Eukaryot. Cell* **2**, 930-936 (2003).
42. Goehring,A.S., Rivers,D.M., & Sprague,G.F., Jr. Urm1ylation: a ubiquitin-like pathway that functions during invasive growth and budding in yeast. **14**, 4329-4341 (2003).
43. Jin,J., Li,X., Gygi,S.P., & Harper,J.W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature* **447**, 1135-1138 (2007).
44. Haas,A.L., Warms,J.V., Hershko,A., & Rose,I.A. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J. Biol. Chem.* **257**, 2543-2548 (1982).

45. Pickart,C.M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503-533 (2001).
46. Tanaka,T., Yeh,E.T., & Kamitani,T. NUB1-mediated targeting of the ubiquitin precursor UbC1 for its C-terminal hydrolysis. *Eur. J. Biochem.* **271**, 972-982 (2004).
47. Finley,D., Bartel,B., & Varshavsky,A. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* **338**, 394-401 (1989).
48. Tanaka,K., Matsumoto,K., & Toh-e A Dual regulation of the expression of the polyubiquitin gene by cyclic AMP and heat shock in yeast. *EMBO J.* **7**, 495-502 (1988).
49. Edelman,M.J. & Kessler,B.M. Ubiquitin and ubiquitin-like specific proteases targeted by infectious pathogens: Emerging patterns and molecular principles. *Biochim. Biophys. Acta* **1782**, 809-816 (2008).
50. Morimoto,M., Nishida,T., Honda,R., & Yasuda,H. Modification of cullin-1 by ubiquitin-like protein Nedd8 enhances the activity of SCF(skp2) toward p27(kip1). *Biochem. Biophys. Res. Commun.* **270**, 1093-1096 (2000).
51. Morimoto,M., Nishida,T., Nagayama,Y., & Yasuda,H. Nedd8-modification of Cul1 is promoted by Roc1 as a Nedd8-E3 ligase and regulates its stability. *Biochem. Biophys. Res. Commun.* **301**, 392-398 (2003).
52. Chen,L. & Chen,J. MDM2-ARF complex regulates p53 sumoylation. *Oncogene* **22**, 5348-5357 (2003).

53. Kwek,S.S., Derry,J., Tyner,A.L., Shen,Z., & Gudkov,A.V. Functional analysis and intracellular localization of p53 modified by SUMO-1. *Oncogene* **20**, 2587-2599 (2001).
54. Minty,A., Dumont,X., Kaghad,M., & Caput,D. Covalent modification of p73alpha by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *J. Biol. Chem.* **275**, 36316-36323 (2000).
55. Bachmair,A., Finley,D., & Varshavsky,A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179-186 (1986).
56. Varshavsky,A., Turner,G., Du,F., & Xie,Y. Felix Hoppe-Seyler Lecture 2000. The ubiquitin system and the N-end rule pathway. *Biol. Chem.* **381**, 779-789 (2000).
57. Wang,K.H., Sauer,R.T., & Baker,T.A. ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes Dev.* **21**, 403-408 (2007).
58. Mogk,A., Schmidt,R., & Bukau,B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol.* **17**, 165-172 (2007).
59. Wang,K.H., Roman-Hernandez,G., Grant,R.A., Sauer,R.T., & Baker,T.A. The molecular basis of N-end rule recognition. *Mol. Cell* **32**, 406-414 (2008).
60. Xie,Y. & Varshavsky,A. Physical association of ubiquitin ligases and the 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A* **97**, 2497-2502 (2000).
61. Shumway,S.D., Maki,M., & Miyamoto,S. The PEST domain of IkappaBalpha is necessary and sufficient for in vitro degradation by mu-calpain. *J. Biol. Chem.* **274**, 30874-30881 (1999).

62. Winston, J.T. *et al.* The SCF $\beta$ -TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in I $\kappa$ B $\alpha$  and  $\beta$ -catenin and stimulates I $\kappa$ B $\alpha$  ubiquitination in vitro. *Genes Dev.* **13**, 270-283 (1999).
63. Zachariae, W. Destruction with a box: substrate recognition by the anaphase-promoting complex. *Mol. Cell* **13**, 2-3 (2004).
64. Lupas, A.N. & Koretke, K.K. Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *J. Struct. Biol.* **141**, 77-83 (2003).
65. Rao, N.N., Gomez-Garcia, M.R., & Kornberg, A. Inorganic polyphosphate: essential for growth and survival. *Annu. Rev. Biochem.* **78**, 605-647 (2009).
66. Kuroda, A., Nomura, K., Takiguchi, N., Kato, J., & Ohtake, H. Inorganic polyphosphate stimulates Lon-mediated proteolysis of nucleoid proteins in *Escherichia coli*. *Cell Mol. Biol. (Noisy. -le-grand)* **52**, 23-29 (2006).
67. Kuroda, A. *et al.* Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli*. *Science* **293**, 705-708 (2001).
68. Docampo, R., de, S.W., Miranda, K., Rohloff, P., & Moreno, S.N. Acidocalcisomes - conserved from bacteria to man. *Nat. Rev. Microbiol.* **3**, 251-261 (2005).
69. Kumble, K.D. & Kornberg, A. Inorganic polyphosphate in mammalian cells and tissues. *J. Biol. Chem.* **270**, 5818-5822 (1995).

70. Scott,P.H., Brunn,G.J., Kohn,A.D., Roth,R.A., & Lawrence,J.C., Jr. Evidence of insulin-stimulated phosphorylation and activation of the mammalian target of rapamycin mediated by a protein kinase B signaling pathway. *Proc. Natl. Acad. Sci. U. S. A* **95**, 7772-7777 (1998).
71. Wang,L., Fraley,C.D., Faridi,J., Kornberg,A., & Roth,R.A. Inorganic polyphosphate stimulates mammalian TOR, a kinase involved in the proliferation of mammary cancer cells. *Proc. Natl. Acad. Sci. U. S. A* **100**, 11249-11254 (2003).
72. Gottesman,S., Roche,E., Zhou,Y., & Sauer,R.T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* **12**, 1338-1347 (1998).
73. Lies,M. & Maurizi,M.R. Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *J. Biol. Chem.* **283**, 22918-22929 (2008).
74. Roche,E.D. & Sauer,R.T. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J* **18**, 4579-4589 (1999).
75. Withey,J.H. & Friedman,D.I. A salvage pathway for protein structures: tmRNA and trans-translation. *Annu. Rev. Microbiol* **57**, 101-123 (2003).
76. Adhikari,A. & Chen,Z.J. Diversity of polyubiquitin chains. *Dev. Cell* **16**, 485-486 (2009).
77. Peng,J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* **21**, 921-926 (2003).



78. Kirisako, T. *et al.* A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J.* **25**, 4877-4887 (2006).
79. Kim, H.T. *et al.* Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol. Chem.* **282**, 17375-17386 (2007).
80. Lallemand-Breitenbach, V. *et al.* Arsenic degrades PML or PML-RAR $\alpha$  through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* **10**, 547-555 (2008).
81. Tatham, M.H. *et al.* RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* **10**, 538-546 (2008).
82. Thrower, J.S., Hoffman, L., Rechsteiner, M., & Pickart, C.M. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94-102 (2000).
83. Pickart, C.M. Ubiquitin in chains. *Trends Biochem. Sci.* **25**, 544-548 (2000).
84. Pickart, C.M. & Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr. Opin. Chem. Biol.* **8**, 610-616 (2004).
85. Carter, P. & Wells, J.A. Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564-568 (1988).
86. Kossiakoff, A.A. & Spencer, S.A. Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. *Biochemistry* **20**, 6462-6474 (1981).

87. Vernet,T. *et al.* Structural and functional roles of asparagine 175 in the cysteine protease papain. *J. Biol. Chem.* **270**, 16645-16652 (1995).
88. Yabuki,Y., Kubota,K., Kojima,M., Inoue,H., & Takahashi,K. Identification of a glutamine residue essential for catalytic activity of aspergilloglutamic peptidase by site-directed mutagenesis. *FEBS Lett.* **569**, 161-164 (2004).
89. Heinemeyer,W., Ramos,P.C., & Dohmen,R.J. The ultimate nanoscale mincer: assembly, structure and active sites of the 20S proteasome core. *Cell Mol. Life Sci.* **61**, 1562-1578 (2004).
90. Dye,B.T. & Schulman,B.A. Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 131-150 (2007).
91. Reyes-Turcu,F.E., Ventii,K.H., & Wilkinson,K.D. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu. Rev. Biochem.* **78**, 363-397 (2009).
92. Amerik,A.Y. & Hochstrasser,M. Mechanism and function of deubiquitinating enzymes. *Biochim. Biophys. Acta* **1695**, 189-207 (2004).
93. Nanao,M.H. *et al.* Crystal structure of human otubain 2. *EMBO Rep.* **5**, 783-788 (2004).
94. Semple,C.A. The comparative proteomics of ubiquitination in mouse. *Genome Res.* **13**, 1389-1394 (2003).
95. Burns,K.E. *et al.* Reconstitution of a new cysteine biosynthetic pathway in *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* **127**, 11602-11603 (2005).

96. Scheel,H., Tomiuk,S., & Hofmann,K. Elucidation of ataxin-3 and ataxin-7 function by integrative bioinformatics. *Hum. Mol. Genet.* **12**, 2845-2852 (2003).
97. Kirisako,T. *et al.* The reversible modification regulates the membrane-binding state of Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole targeting pathway. *J. Cell Biol.* **151**, 263-276 (2000).
98. Ha,B.H. & Kim,E.E. Structures of proteases for ubiquitin and ubiquitin-like modifiers. *BMB. Rep.* **41**, 435-443 (2008).
99. Pickart,C.M. & Rose,I.A. Ubiquitin carboxyl-terminal hydrolase acts on ubiquitin carboxyl-terminal amides. *J. Biol. Chem.* **260**, 7903-7910 (1985).
100. Geng,J. & Klionsky,D.J. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. 'Protein modifications: beyond the usual suspects' review series. *EMBO Rep.* **9**, 859-864 (2008).
101. Beyaert,R., Heyninck,K., & Van,H.S. A20 and A20-binding proteins as cellular inhibitors of nuclear factor-kappa B-dependent gene expression and apoptosis. *Biochem. Pharmacol.* **60**, 1143-1151 (2000).
102. Soares,L. *et al.* Two isoforms of otubain 1 regulate T cell anergy via GRAIL. *Nat. Immunol.* **5**, 45-54 (2004).
103. Winborn,B.J. *et al.* The deubiquitinating enzyme ataxin-3, a polyglutamine disease protein, edits Lys63 linkages in mixed linkage ubiquitin chains. *J. Biol. Chem.* **283**, 26436-26443 (2008).

104. Doss-Pepe,E.W., Stenroos,E.S., Johnson,W.G., & Madura,K. Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis. *Mol. Cell Biol.* **23**, 6469-6483 (2003).
105. Bornstein,G., Ganoth,D., & Hershko,A. Regulation of neddylation and deneddylation of cullin1 in SCFSkp2 ubiquitin ligase by F-box protein and substrate. *Proc. Natl. Acad. Sci. U. S. A* **103**, 11515-11520 (2006).
106. Harty,R.N., Pitha,P.M., & Okumura,A. Antiviral Activity of Innate Immune Protein ISG15. *J. Innate. Immun.* **1**, 397-404 (2009).
107. Roden,J., Eardley,L., Hotson,A., Cao,Y., & Mudgett,M.B. Characterization of the Xanthomonas AvrXv4 effector, a SUMO protease translocated into plant cells. *Mol. Plant Microbe Interact.* **17**, 633-643 (2004).
108. Misaghi,S. *et al.* Chlamydia trachomatis-derived deubiquitinating enzymes in mammalian cells during infection. *Mol. Microbiol.* **61**, 142-150 (2006).
109. Gong,L., Kamitani,T., Millas,S., & Yeh,E.T. Identification of a novel isopeptidase with dual specificity for ubiquitin- and NEDD8-conjugated proteins. *J. Biol. Chem.* **275**, 14212-14216 (2000).
110. Sowa,M.E., Bennett,E.J., Gygi,S.P., & Harper,J.W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389-403 (2009).
111. Heyninck,K. & Beyaert,R. A20 inhibits NF-kappaB activation by dual ubiquitin-editing functions. *Trends Biochem. Sci.* **30**, 1-4 (2005).

112. Makarova,K.S., Aravind,L., & Koonin,E.V. A novel superfamily of predicted cysteine proteases from eukaryotes, viruses and *Chlamydia pneumoniae*. *Trends Biochem. Sci.* **25**, 50-52 (2000).
113. Opiari,A.W., Jr., Boguski,M.S., & Dixit,V.M. The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein. *J. Biol. Chem.* **265**, 14705-14708 (1990).
114. Evans,P.C. *et al.* Zinc-finger protein A20, a regulator of inflammation and cell survival, has de-ubiquitinating activity. *Biochem. J.* **378**, 727-734 (2004).
115. Wertz,I.E. *et al.* De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature* **430**, 694-699 (2004).
116. Boone,D.L. *et al.* The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat. Immunol.* **5**, 1052-1060 (2004).
117. Shen,L.N. *et al.* Structural basis of NEDD8 ubiquitin discrimination by the deNEDDylating enzyme NEDP1. *EMBO J* **24**, 1341-1351 (2005).
118. Johnston,S.C., Larsen,C.N., Cook,W.J., Wilkinson,K.D., & Hill,C.P. Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 Å resolution. *EMBO J.* **16**, 3787-3796 (1997).
119. Misaghi,S. *et al.* Structure of the ubiquitin hydrolase UCH-L3 complexed with a suicide substrate. *J. Biol. Chem.* **280**, 1512-1520 (2005).
120. Popp,M.W., Artavanis-Tsakonas,K., & Ploegh,H.L. Substrate filtering by the active site crossover loop in UCHL3 revealed by sortagging and gain-of-function mutations. *J. Biol. Chem.* **284**, 3593-3602 (2009).

121. Saito,K. *et al.* The CAP-Gly domain of CYLD associates with the proline-rich sequence in NEMO/IKKgamma. *Structure*. **12**, 1719-1728 (2004).
122. Reyes-Turcu,F.E. *et al.* The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **124**, 1197-1208 (2006).
123. Dang,L.C., Melandri,F.D., & Stein,R.L. Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. *Biochemistry* **37**, 1868-1879 (1998).
124. Crosas,B. *et al.* Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. *Cell* **127**, 1401-1413 (2006).
125. Yao,T. *et al.* Distinct modes of regulation of the Uch37 deubiquitinating enzyme in the proteasome and in the Ino80 chromatin-remodeling complex. *Mol. Cell* **31**, 909-917 (2008).
126. Hutti,J.E. *et al.* IkappaB kinase beta phosphorylates the K63 deubiquitinase A20 to cause feedback inhibition of the NF-kappaB pathway. *Mol. Cell Biol.* **27**, 7451-7461 (2007).
127. Reiley,W., Zhang,M., Wu,X., Granger,E., & Sun,S.C. Regulation of the deubiquitinating enzyme CYLD by IkappaB kinase gamma-dependent phosphorylation. *Mol. Cell Biol.* **25**, 3886-3895 (2005).
128. Todi,S.V. *et al.* Ubiquitination directly enhances activity of the deubiquitinating enzyme ataxin-3. *EMBO J.* **28**, 372-382 (2009).

129. Whitby,F.G., Xia,G., Pickart,C.M., & Hill,C.P. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J. Biol. Chem.* **273**, 34983-34991 (1998).
130. Komander,D., Clague,M.J., & Urbe,S. Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550-563 (2009).
131. Striebel,F., Kress,W., & Weber-Ban,E. Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. *Curr. Opin. Struct. Biol.* **19**, 209-217 (2009).
132. Maupin-Furlow,J.A., Wilson,H.L., Kaczowka,S.J., & Ou,M.S. Proteasomes in the archaea: from structure to function. *Front Biosci.* **5**, D837-D865 (2000).
133. Gottesman,S. Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell Dev. Biol.* **19**, 565-587 (2003).
134. Mueller,C.A., Broz,P., & Cornelis,G.R. The type III secretion system tip complex and translocon. *Mol. Microbiol.* **68**, 1085-1095 (2008).
135. Mecsas,J.J. & Strauss,E.J. Molecular mechanisms of bacterial virulence: type III secretion and pathogenicity islands. *Emerg. Infect. Dis.* **2**, 270-288 (1996).
136. Cornelis,G.R. The type III secretion injectisome. *Nat. Rev. Microbiol.* **4**, 811-825 (2006).
137. Janjusevic,R., Abramovitch,R.B., Martin,G.B., & Stebbins,C.E. A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* **311**, 222-226 (2006).

138. Butt,T.R. *et al.* Ubiquitin fusion augments the yield of cloned gene products in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **86**, 2540-2544 (1989).
139. Gonen,H. *et al.* Protein synthesis elongation factor EF-1 alpha is essential for ubiquitin-dependent degradation of certain N alpha-acetylated proteins and may be substituted for by the bacterial elongation factor EF-Tu. *Proc. Natl. Acad. Sci. U. S. A* **91**, 7648-7652 (1994).
140. Berchtold,H. *et al.* Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **365**, 126-132 (1993).
141. Bourne,H.R., Sanders,D.A., & McCormick,F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348**, 125-132 (1990).
142. Bourne,H.R., Sanders,D.A., & McCormick,F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117-127 (1991).
143. Tsirigotis,M. *et al.* Analysis of ubiquitination in vivo using a transgenic mouse model. *Biotechniques* **31**, 120-6, 128, 130 (2001).
144. Gilchrist,C.A., Gray,D.A., & Baker,R.T. A ubiquitin-specific protease that efficiently cleaves the ubiquitin-proline bond. *J. Biol. Chem.* **272**, 32280-32285 (1997).
145. Fath,M.J. & Kolter,R. ABC transporters: bacterial exporters. *Microbiol. Rev.* **57**, 995-1017 (1993).
146. Kerr,I.D., Reynolds,E.D., & Cove,J.H. ABC proteins and antibiotic drug resistance: is it all about transport? *Biochem. Soc. Trans.* **33**, 1000-1002 (2005).



147. McDevitt, C.A. *et al.* Is ATP binding responsible for initiating drug translocation by the multidrug transporter ABCG2? *FEBS J.* **275**, 4354-4362 (2008).
148. Butt, T.R., Khan, M.I., Marsh, J., Ecker, D.J., & Crooke, S.T. Ubiquitin-metallothionein fusion protein expression in yeast. A genetic approach for analysis of ubiquitin functions. *J. Biol. Chem.* **263**, 16364-16371 (1988).
149. Giglione, C., Boularot, A., & Meinnel, T. Protein N-terminal methionine excision. *Cell Mol. Life Sci.* **61**, 1455-1474 (2004).
150. Geoghegan, K.F. *et al.* Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **267**, 169-184 (1999).
151. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A* **97**, 6640-6645 (2000).
152. Rudolph, M.J., Wuebbens, M.M., Rajagopalan, K.V., & Schindelin, H. Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. *Nat. Struct. Biol.* **8**, 42-46 (2001).
153. Wang, C., Xi, J., Begley, T.P., & Nicholson, L.K. Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. *Nat. Struct. Biol.* **8**, 47-51 (2001).
154. Leimkuhler, S., Wuebbens, M.M., & Rajagopalan, K.V. Characterization of *Escherichia coli* MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the

- molybdenum cofactor. *J. Biol. Chem.* **276**, 34695-34701 (2001).
155. Taylor, S.V. *et al.* Thiamin biosynthesis in *Escherichia coli*. Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation. *J. Biol. Chem.* **273**, 16555-16560 (1998).
  156. Lubelski, J., Mazurkiewicz, P., van, M.R., Konings, W.N., & Driessen, A.J. *ydaG* and *ydbA* of *Lactococcus lactis* encode a heterodimeric ATP-binding cassette-type multidrug transporter. *J. Biol. Chem.* **279**, 34449-34455 (2004).
  157. Harper, J.W. & Schulman, B.A. Structural complexity in ubiquitin recognition. *Cell* **124**, 1133-1136 (2006).
  158. Wilkinson, K.D. *et al.* Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* **34**, 14535-14546 (1995).
  159. Wang, H. *et al.* Analysis of nondegradative protein ubiquitylation with a monoclonal antibody specific for lysine-63-linked polyubiquitin. *Proc. Natl. Acad. Sci. U. S. A* **105**, 20197-20202 (2008).
  160. Huang, J. *et al.* ZNF216 Is an A20-like and I $\kappa$ B kinase  $\gamma$ -interacting inhibitor of NF $\kappa$ B activation. *J. Biol. Chem.* **279**, 16847-16853 (2004).
  161. Heng, A.E. *et al.* Coordinate expression of the 19S regulatory complex and evidence for ubiquitin-dependent telethonin degradation in the unloaded soleus muscle. *Int. J. Biochem. Cell Biol.* **40**, 2544-2552 (2008).
  162. Papanikou, E., Karamanou, S., & Economou, A. Bacterial protein secretion through the translocase

- nanomachine. *Nat. Rev. Microbiol.* **5**, 839-851 (2007).
163. Blango, M.G. & Mulvey, M.A. Bacterial landlines: contact-dependent signaling in bacterial populations. *Curr. Opin. Microbiol.* **12**, 177-181 (2009).
164. Slechta, E.S. & Mulvey, M.A. Contact-dependent inhibition: bacterial brakes and secret handshakes. *Trends Microbiol.* **14**, 58-60 (2006).
165. Reyes-Turcu, F.E., Shanks, J.R., Komander, D., & Wilkinson, K.D. Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T. *J. Biol. Chem.* **283**, 19581-19592 (2008).
166. Mendoza, H.M. *et al.* NEDP1, a highly conserved cysteine protease that deNEDDylates Cullins. *J. Biol. Chem.* **278**, 25637-25643 (2003).
167. Johnston, S.C., Riddle, S.M., Cohen, R.E., & Hill, C.P. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J.* **18**, 3877-3887 (1999).
168. Mossessova, E. & Lima, C.D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* **5**, 865-876 (2000).
169. Balakirev, M.Y., Jaquinod, M., Haas, A.L., & Chroboczek, J. Deubiquitinating function of adenovirus proteinase. *J. Virol.* **76**, 6323-6331 (2002).
170. Kelly, D. *et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat. Immunol.* **5**, 104-112 (2004).

171. Tien, M.T. *et al.* Anti-inflammatory effect of *Lactobacillus casei* on *Shigella*-infected human intestinal epithelial cells. *J. Immunol.* **176**, 1228-1237 (2006).
172. Iwai, K. & Tokunaga, F. Linear polyubiquitination: a new regulator of NF-kappaB activation. *EMBO Rep.* **10**, 706-713 (2009).
173. Karagiannis, J. & Young, P.G. Intracellular pH homeostasis during cell-cycle progression and growth state transition in *Schizosaccharomyces pombe*. *J. Cell Sci.* **114**, 2929-2941 (2001).